

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 3, 2005, 17:11:04 ; Search time 7197 Seconds  
(without alignments)

11847.153 Million cell updates/sec

Title: US-09-544-776-1

Perfect score: 2240

Sequence: 1 cgtccacacagtaggtccct.....taaaaaaaaaaaaaaaaaaa 2240

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 159776

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: gb\_est1:\*

2: gb\_est2:\*

3: gb\_hic:\*

4: gb\_est3:\*

5: gb\_est4:\*

6: gb\_est5:\*

7: gb\_est6:\*

8: gb\_g881:\*

9: gb\_g882:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	41.8	1.9	46	1 AA995439	AA995439 0884f01.8
2	36	1.6	50	1 AU104978	AU104978 AU104978
3	36	1.6	50	1 AU104979	AU104979 AU104979
4	36	1.6	50	1 AU104981	AU104981 AU104981
5	36	1.6	50	1 AU104985	AU104985 AU104985
6	34.4	1.5	50	1 AU104974	AU104974 AU104974
7	34.4	1.5	50	1 AU104977	AU104977 AU104977
8	34.4	1.5	50	1 AU104980	AU104980 AU104980
9	34	1.5	50	1 AU104976	AU104976 AU104976
10	33.8	1.5	37	8 AZ761912	AZ761912 1M0556D02
11	33	1.5	50	1 AU104983	AU104983 AU104983
C 12	32.2	1.4	45	8 AZ843544	AZ843544 2M0142022
13	32	1.4	50	1 AU104984	AU104984 AU104984
14	32	1.4	50	1 AU104986	AU104986 AU104986
C 15	30.8	1.4	50	9 BX976589	BX976589 Forward s
16	30.2	1.3	43	8 AZ419099	AZ419099 1M0195N23
17	29.6	1.3	37	8 AZ392980	AZ392980 1M0155P13
18	29.6	1.3	37	8 AZ623276	AZ623276 1M0460M16
19	29.6	1.3	48	8 AZ443723	AZ443723 1M0238D11
20	29.2	1.3	43	7 H41505	H41505 yp68h12.s1
21	29.2	1.3	50	1 A1252059	A1252059 qv39f04.x
C 22	28.8	1.3	49	8 AZ840678	AZ840678 2M0138O16
23	28.6	1.3	36	8 AZ479840	AZ479840 1M0300P16
24	28.6	1.3	49	2 BF789094	BF789094 602104956

#### ALIGNMENTS

RESULT 1  
AA995439/c

LOCUS  
DEFINITION

AA995439

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

28.2 1.3 50 9 CG731182  
27.2 1.2 32 8 AZ792853  
27 1.2 50 1 AU103294  
26.6 1.2 33 8 AZ435186  
26.6 1.2 50 6 CD577044  
26.6 1.2 50 7 CR430378  
26.6 1.2 50 8 AZ627160  
26.4 1.2 47 2 BF107886  
26.4 1.2 49 1 AA145482  
26.4 1.2 50 1 AA590944  
26.2 1.2 44 8 AZ968544  
26.2 1.2 49 8 AZ407440  
26 1.2 50 4 BI491464  
26 1.2 50 7 CV304274  
26 1.2 50 8 AZ358097  
26 1.2 50 8 AZ816605  
26 1.2 50 8 AZ861629  
26 1.2 50 8 AZ942114  
26 1.2 50 8 AZ967786  
26 1.2 50 9 CR003825  
26 1.2 50 9 CR048038

AA995439 46 bp mRNA linear EST 10-NOV-1998  
similar to TR:Q90637 Q90637 CHS-REX-S.; mRNA sequence.  
AA995439 GI:3181928  
EST.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.  
Emmert-Buck, M.D., Ph.D.  
CDNA Library Preparation: Stratagene, Inc., David B. Krizman,  
Ph.D.  
CDNA Library Arraying: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
www-bio.llnl.gov/bbrp/image/image.html  
Trace considered overall poor quality  
Insert Length: 1604 Std Error: 0.00  
Seq primer: -40m13 fwd. RT from Amersham  
High quality sequence stop: 1.  
Location/Qualifiers  
1. .46  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:1612057"  
/tissue\_type="germ cell tumor"  
/lab\_host="SOLR (kanamycin resistant)"  
/clone\_lib="NCI CGAP GC2"  
/notes="Vector: Bluescript SK-; Site 1: EcoRI; Site 2:  
XhoI; Cloned unidirectionally. Primer: Oligo dT. Bulk  
germ cell tumor. 5' adaptor sequence: 5' GAATTCGACGACG 3'  
3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'  
Average insert size: 1.2 kb."



```

Query Match      1.6%; Score 36; DB 1; Length 50;
Best Local Similarity 100.0%; Pred. No. 3.9e+03;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTCACACAGTAGTCCCTCGGCTCAGTCGGGCCCA 36
    |||
Db 15 CGTCACACAGTAGTCCCTCGGCTCAGTCGGGCCCA 50

RESULT 5
LOCUS AU104985
DEFINITION AU104985 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
          EST 28-JAN-2004
ACCESSION HSI01927, mRNA sequence.
VERSION AU104985
KEYWORDS AU104985.1 GI:13554506
SOURCE EST.
ORGANISM Homo sapiens (human)
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J.,
          Hata,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K.,
          Sakaki,Y., Yoshitomo-Nakagawa,K., Maruyama,A. and Sugano,S.
TITLE Diverse transcriptional initiation revealed by fine, large-scale
mapping of mRNA start sites
JOURNAL EMBO Rep. 2 (5), 388-393 (2001)
MEDLINE 21270072
PUBMED 11375929
COMMENT Contact: Yutaka Suzuki
          Department of Virology
          Institute of Medical Science, University of Tokyo
          4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
          Email: ysuzuki@ims.u-tokyo.ac.jp
          Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Maruyama,A. and
          Sugano,S. Construction and characterization of a full
          length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2),
          149-156 (1997).

FEATURES
source
    1..50
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="CAE01363"
    /clone_lib="Sugano Homo sapiens cDNA library"

ORIGIN
Query Match      1.5%; Score 34.4; DB 1; Length 50;
Best Local Similarity 97.2%; Pred. No. 8.8e+03;
Matches 35; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGTCACACAGTAGTCCCTCGGCTCAGTCGGGCCCA 36
    |||
Db 15 CGTCACACAGTAGTCCCTCGGCTCAGTCGGGCCCA 50

RESULT 7
LOCUS AU104977
DEFINITION AU104977 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
          EST 28-JAN-2004
ACCESSION CAS03274, mRNA sequence.
VERSION AU104977
KEYWORDS AU104977.1 GI:13554498
SOURCE EST.
ORGANISM Homo sapiens (human)
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J.,
          Hata,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K.,
          Sakaki,Y., Nakamura,Y., Suyama,A. and Sugano,S.
TITLE Diverse transcriptional initiation revealed by fine, large-scale
mapping of mRNA start sites
JOURNAL EMBO Rep. 2 (5), 388-393 (2001)
MEDLINE 21270072
PUBMED 11375929
COMMENT Contact: Yutaka Suzuki
          Department of Virology
          Institute of Medical Science, University of Tokyo
          4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
          Email: ysuzuki@ims.u-tokyo.ac.jp
          Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Maruyama,A. and
          Sugano,S. Construction and characterization of a full
          length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2),
          149-156 (1997).

FEATURES
source
    1..50
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="HSI01927"
    /clone_lib="Sugano Homo sapiens cDNA library"

ORIGIN
Query Match      1.6%; Score 36; DB 1; Length 50;
Best Local Similarity 100.0%; Pred. No. 3.9e+03;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTCACACAGTAGTCCCTCGGCTCAGTCGGGCCCA 36
    |||
Db 15 CGTCACACAGTAGTCCCTCGGCTCAGTCGGGCCCA 50

RESULT 6
LOCUS AU104974
DEFINITION AU104974 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
          EST 28-JAN-2004
ACCESSION CAE01363, mRNA sequence.
VERSION AU104974
KEYWORDS AU104974.1 GI:13554495
SOURCE EST.
ORGANISM Homo sapiens (human)
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J.,
          Hata,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K.,
          Sakaki,Y., Nakamura,Y., Suyama,A. and Sugano,S.

```

```

TITLE Diverse transcriptional initiation revealed by fine, large-scale
mapping of mRNA start sites
JOURNAL EMBO Rep. 2 (5), 388-393 (2001)
MEDLINE 21270072
PUBMED 11375929
COMMENT Contact: Yutaka Suzuki
          Department of Virology
          Institute of Medical Science, University of Tokyo
          4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
          Email: ysuzuki@ims.u-tokyo.ac.jp
          Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Maruyama,A. and
          Sugano,S. Construction and characterization of a full
          length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2),
          149-156 (1997).

FEATURES
source
    1..50
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="CAE01363"
    /clone_lib="Sugano Homo sapiens cDNA library"

ORIGIN
Query Match      1.5%; Score 34.4; DB 1; Length 50;
Best Local Similarity 97.2%; Pred. No. 8.8e+03;
Matches 35; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGTCACACAGTAGTCCCTCGGCTCAGTCGGGCCCA 36
    |||
Db 15 CGTCACACAGTAGTCCCTCGGCTCAGTCGGGCCCA 50

RESULT 7
LOCUS AU104977
DEFINITION AU104977 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
          EST 28-JAN-2004
ACCESSION CAS03274, mRNA sequence.
VERSION AU104977
KEYWORDS AU104977.1 GI:13554498
SOURCE EST.
ORGANISM Homo sapiens (human)
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J.,
          Hata,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K.,
          Sakaki,Y., Nakamura,Y., Suyama,A. and Sugano,S.
TITLE Diverse transcriptional initiation revealed by fine, large-scale
mapping of mRNA start sites
JOURNAL EMBO Rep. 2 (5), 388-393 (2001)
MEDLINE 21270072
PUBMED 11375929
COMMENT Contact: Yutaka Suzuki
          Department of Virology
          Institute of Medical Science, University of Tokyo
          4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
          Email: ysuzuki@ims.u-tokyo.ac.jp
          Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Maruyama,A. and
          Sugano,S. Construction and characterization of a full
          length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2),
          149-156 (1997).

FEATURES
source
    1..50
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="CAS03274"
    /clone_lib="Sugano Homo sapiens cDNA library"

ORIGIN
Query Match      1.5%; Score 34.4; DB 1; Length 50;
Best Local Similarity 97.2%; Pred. No. 8.8e+03;

```





/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (GI|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

## ORIGIN

Query Match 1.5%; Score 33.8; DB 8; Length 37;  
Best Local Similarity 94.6%; Pred. No. 1.1e+04;  
Matches 35; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 237 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 273  
Db 1 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 37

## RESULT 11

AU104983 50 bp mRNA linear EST 28-JAN-2004  
LOCUS AU104983 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone  
DEFINITION HRC11726, mRNA sequence.

ACCESSION AU104983.1 GI:13554504  
VERSION EST.  
KEYWORDS  
SOURCE Homo sapiens (human)

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

1 (bases 1 to 50)  
Suzuki, Y., Taira, H., Tsunoda, T., Mizushima-Sugano, J., Sese, J.,  
Hata, H., Oka, T., Isogai, T., Tanaka, T., Morishita, S., Okubo, K.,  
Sakaki, Y., Nakamura, Y., Suyama, A. and Sugano, S.

Diverse transcriptional initiation revealed by fine, large-scale

mapping of mRNA start sites  
EMBO Rep. 2 (5), 388-393 (2001)

## MEDLINE

## PUBMED

## COMMENT

Contact: Yutaka Suzuki

Institute of Medical Science, University of Tokyo

4-6-1, Shirokanedai, Minato-ku, Tokyo 108-8639, Japan

Email: [shirokai@ims.u-tokyo.ac.jp](mailto:shirokai@ims.u-tokyo.ac.jp)

Suzuki, Y., Yoshitomo-Nakagawa, K., Maruyama, K., Suyama, A. and

Sugano, S. Construction and characterization of a full

length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2),

149-156 (1997).

## FEATURES

## source

1. 50  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="HRC11726"  
/clone\_lib="Sugano Homo sapiens cDNA library"

## ORIGIN

Query Match 1.5%; Score 33; DB 1; Length 50;  
Best Local Similarity 100.0%; Pred. No. 1.8e+04;  
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

## QY

5 ACCACAGTAGGTCCTCGGCTCAGTCGCGCCAG 37  
Db 18 ACCACAGTAGGTCCTCGGCTCAGTCGCGCCAG 50

## RESULT 12

## LOCUS

## DEFINITION

AZ843544 45 bp DNA linear GSS 20-FEB-2001  
2M0142022F Mouse 10kb plasmid UUGC1M library Mus musculus genomic

clone UUGC2M0142022 F, genomic survey sequence.

## ACCESSION

AZ843544

## KEYWORDS

GSS.

## SOURCE

Mus musculus (house mouse)

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 45)

Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,

Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T.,

Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von

Niederhauser, A. and Wright, D. Weiss, R.

Mouse whole genome scaffolding with paired end reads from 10kb

plasmid inserts

Unpublished (2000)

Contact: Robert B. Weiss

University of Utah Genome Center

University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT

84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: [ddunn@genetics.utah.edu](mailto:ddunn@genetics.utah.edu)

Insert Length: 10000 Std Error: 0.00

Plate: 0142 row: 0 column: 22

Seq primer: CGTTGTAAACGACGCCAGT

Class: plasmid ends

High quality sequence stop: 45.

Location/Qualifiers

1. 45

/organism="Mus musculus"

/mol\_type="genomic DNA"

/strain="C57BL/6J"

/db\_xref="taxon:10090"

/clone="UUGC2M0142022"

/sex="Male"

/lab\_host="E. Coli strain XL10-Gold, Tl-resistant, F-"

/clone\_lib="Mouse 10kb plasmid UUGC1M library"

/note="Vector: PWD42nv; Purified genomic DNA from M.

musculus C57BL/6J (male) was obtained from the Jackson

Laboratory Mouse DNA Resource

(<http://www.jax.org/resources/documents/dnares/>). The DNA

was hydrodynamically sheared by repeated passage through a

0.005 inch orifice at constant velocity. The sheared DNA

was blunt end-repaired with T4 DNA polymerase and T4

polynucleotide kinase. Adaptor oligonucleotides were

ligated to the blunt ends in high molar excess. The

adaptor DNA was purified and size-selected for a 9.5 to

10.5 kb range using preparative agarose gel

electrophoresis. Vector DNA was prepared from a derivative

of pWD42 (GI|4732114|gb|AF129072.1), a copy-number

inducible derivative of plasmid R1. The vector was ligated

with adaptors complementary to the insert adaptors and

purified. The sheared, adaptor mouse DNA was annealed to

adaptor vector DNA, and transformed into

chemically-competent E. coli XL10-Gold (Stratagene) cells

and selected for ampicillin resistance."

## ORIGIN

Query Match 1.4%; Score 32.2; DB 8; Length 45;  
Best Local Similarity 91.9%; Pred. No. 2.6e+04;  
Matches 34; Conservative 0; Mismatches 3; Indels 0; Gaps 0;



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 3, 2005, 17:13:52 ; Search time 4324 Seconds  
(without alignments)  
4179.879 Million cell updates/sec

Title: US-09-544-776-2

Perfect score: 1910  
Sequence: 1 MEDLDQSLVSSSDSPRPQ.....VKDAMAKIAQIKPLKRAE 373

Scoring table: BLOSUM62

Qgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DRV=xlp  
-Q/cgn2\_1/USPTO\_spool\_p/US09544776/runat\_03082005\_181345\_28204/app\_query.fasta\_1.519  
-DB=GenEmbl -QFWT=fastcap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFWT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09544776 @CGN 1.1 5600 @runat\_03082005\_181345\_28204 -NCPU=6 -ICPU=3  
-NO MAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DLEXT=7

Database :

GenEmbl.\*

1: gb.ba.\*

2: gb.htg.\*

3: gb.in.\*

4: gb.om.\*

5: gb.ov.\*

6: gb.pat.\*

7: gb.ph.\*

8: gb.pl.\*

9: gb.pr.\*

10: gb.ro.\*

11: gb.sts.\*

12: gb.sy.\*

13: gb.un.\*

14: gb.vi.\*

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1879	98.4	1122	6	BD249448
2	1879	98.4	1122	9	HSA251384
3	1879	98.4	1485	9	BC010737 Homo sapi
4	1879	98.4	1610	6	BD231889

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

5	1879	98.4	1619	9	AK129806
6	1879	98.4	1654	9	BC012619
7	1879	98.4	1668	9	BC026788 Homo sapi
8	1879	98.4	1709	9	AB040463
9	1879	98.4	1728	9	BC068991
10	1879	98.4	2235	9	AF148538
11	1879	98.4	2276	9	AF132047
12	1879	98.4	2332	9	AF102277
13	1876	98.2	1525	9	AK130812
14	1875	98.2	1599	6	CQ769577
15	1871	98.0	2052	9	AB015639
16	1862	97.5	1694	6	CQ783030
17	1862	97.5	1694	6	BD127437
18	1862	97.5	1694	9	AK075039
19	1859	97.4	1784	9	BC016165
20	1859	97.4	2389	9	AY102278
21	1717	89.9	1466	9	BC071848
22	1584.5	83.0	2156	10	AY164740
23	1584.5	83.0	2156	10	RN0242962
24	1584.5	83.0	2410	10	BC070879
25	1568.5	82.1	2256	10	AF132046
26	1566	82.0	2782	6	AX700396
27	1566	82.0	2782	10	AY164741
28	1566	82.0	2782	10	AF132045
29	1560.5	81.7	2209	10	AY102281
30	1550	81.2	1276	10	AY114153
31	1542	80.7	2266	10	AY102282
32	1459.5	76.4	3576	6	AX766050
33	1459.5	76.4	3579	6	BD249446
34	1459.5	76.4	3579	9	HSA251383
35	1459.5	76.4	3919	6	CQ829486
36	1459.5	76.4	4053	6	AX195249
37	1459.5	76.4	4053	9	AB020693
38	1459.5	76.4	4166	9	AB040462
39	1459.5	76.4	4832	9	AF148537
40	1459.5	76.4	4789	6	CQ874017
41	1459.5	76.4	4789	9	AY102279
42	1437	75.2	4093	6	BD270070
43	1431	74.9	4822	6	AR220865
44	1258.5	65.9	2934	6	CQ828004
45	1173	61.4	3489	6	AX766046

ALIGNMENTS

RESULT 1	BD249448	1122 bp	DNA	linear	PAT 17-JUL-2003
LOCUS	BD249448	Protein similar to neuroendocrine-specific protein, and encoding			
DEFINITION	BD249448	cDNA.			
ACCESSION	BD249448				
VERSION	BD249448.1	GI:33059218			
KEYWORDS	JP 2002522016-A/3.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	1	(bases 1 to 1122)			
AUTHORS	Michalovich, D. and Prinjha, R.K.				
TITLE	Protein similar to neuroendocrine-specific protein, and encoding				
JOURNAL	Patent: JP 2002522016-A 3 23-JUL-2002;				
COMMENT	SMITHKLINE BEECHAM PLC				
	OS Homo sapiens (human)				
	PN JP 2002522016-A/3				
	PD 23-JUL-2002				
	PF 21-JUL-1999	JP 2000561310			
	PR 22-JUL-1998	GB 9816024.5, 19-JUL-1999	GB	9916898.1	PI
	DAVID MICHALOVICH, RABINDER KUMAR PRINJHA				
	PC				
	C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/				
	10,				
	PC C12P21/02, C12Q1/02, C12Q1/68, G01N33/15, G01N33/50, G01N33/53, PC				
	G01N33/566//				

PC	C12P21/08.C12N15/00.C12N5/00	
CC	Protein similar to neuroendocrine-specific protein, and CC	
FH	Key	Location/Qualifiers
FT	source	1..1122
FT	source	/organism="Homo sapiens (human)"
FEATURES	Location/Qualifiers	
source	1..1122	
	/organism="Homo sapiens"	
	/mol_type="genomic DNA"	
	/db_xref="taxon:9606"	
ORIGIN		
Alignment Scores:	2.97e-67	Length: 1122
Pred. No.:	1879.00	Matches: 370
Score:	99.20%	Conservative: 0
Percent Similarity:	99.20%	Mismatches: 3
Best Local Similarity:	98.38%	Indels: 0
Query Match:	6	Gaps: 0
DB:		
US-09-544-776-2 (1-373) x BD249448 (1-1122)		
QY	1 MetGluAspLeuAspGlnSerProLeuValSerSerAspSerProProArgProGln 20	
DB	1 ATGGAAGACTGGACCAAGTCTCTCTGTCTCGTCTCGACACCCACCCGGCCGAG 60	
QY	21 ProAlaPheLysTyrGlnPheValArgGluProGluAspGluGluGluGlu 40	
DB	61 CCGCGCTTCAAGTACCAGTTCGTGTGGAGAGCCCGAGGAGGAGGAGGAG 120	
QY	41 GluGluAspGluAspGluAspLeuGluGluLeuGluValLeuGluArgLysProAla 60	
DB	121 GAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 180	
QY	61 AlaGlyLeuSerAlaAlaProValProThrAlaProAlaAlaGlyAlaProLeuMetAsp 80	
DB	181 GCCGGGCTGTCCGGCGCCAGTGCACCGCCCTCGCCGCGCGCGCTGATGGAC 240	
QY	81 PheGlyAsnAspPheValProProAlaProArgGlyPheLeuProAlaAlaProProVal 100	
DB	241 TTCGGAATACATCTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300	
QY	101 AlaProGluArgGlnProSerTrpAspProSerProValSerSerThrValProAlaPro 120	
DB	301 GCCCGGAGCGGAGCGGCTTGGGACCGAGCGCGCGCGCGCGCGCGCGCGCG 360	
QY	121 SerPheLeuSerAlaAlaAlaValSerProSerLysLeuProGluAspGluProPro 140	
DB	361 TCCCGCGCTGTCTGCTCGCGAGTCTCGCCCTCCAGCTCCCTGAGGACGAGCCTCG 420	
QY	141 AlaArgProProProProProAlaSerValSerProGlnAlaGluProValTrpThr 160	
DB	421 GCCCGGCGCTCCCGCTCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480	
QY	161 ProProAlaProAlaProAlaProProSerThrProAlaAlaProLysArgArgGly 180	
DB	481 CCGCCAGCGCGGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 540	
QY	181 SerSerGlySerValValValAspLeuLeuTyrTrpArgAspLysLysThrGlyVal 200	
DB	541 TCCTCGGCTCAGTGTGTGACCTCTCTACTCGGAGAGACATTAAAGAACTGGAGTG 600	
QY	201 ValPheGlyAlaSerLeuPheLeuLeuLeuSerLeuThrValPheSerLeuValSerVal 220	
DB	601 GTGTTTGGTCCACGCTATTCTCTGCTGCTTTCATTGACAGTATTCAGCATTTGACGCTA 660	
QY	221 ThrAlaTyrLeuAlaLeuAlaLeuSerValThrIleSerProArgLysTyrLysGly 240	
DB	661 ACAGCCTACATTGCCCTTGGCCCTGCTCTCTGACCATCAGCTTTAGGATATACAGGGT 720	
QY	241 ValIleGlnAlaIleGlnLysSerAspGluGlyHisProPheArgAlaTyrLeuGluSer 260	

Db	721 GTGATCCAAAGCTATCCAGAAATCAGATGAAGCCACCCATTGAGGCATATCTGGAATCT 780	
QY	261 GluValAlaIleSerGluLeuValGlnLysTyrSerAsnSerAlaLeuGlyHisVal 280	
Db	781 GAAGTTGCTATATCTGAGGAGTTGGTTTCAGAAAGTACAGTAATCTCTCTTGGTCAATGTG 840	
QY	281 AsnCysThrIleLysGluLeuArgLeuPheLeuValAspAspLeuValAspSerLeu 300	
Db	841 AACTGCACGATAAAGAACTCAGGGCCCTCTCTTCTAGTTGATGATTTAGTTGATCTCTG 900	
QY	301 LysPheAlaValLeuMetTrpValPheThrTyrValGlyValAlaLeuPheAsnGlyLeuThr 320	
Db	901 AAGTTTGCAGTCTGATGTGGGTATTTACCTATGTTGGTGCCTGTTTAAATGCTCTGACA 960	
QY	321 LeuLeuIleLeuAlaLeuIleSerLeuPheSerValProValIleTyrGluArgHisGln 340	
Db	961 CTACTGATTTTGGCTCTCATTTTCACTTCTTCACTGTTTCTTATTTATGAACGGCATCAG 1020	
QY	341 AlaGlnIleAspHisTyrLeuGlyLeuAlaAsnLysAsnValLysAspAlaMetAlaLys 360	
Db	1021 GCACAGATAGATCATTAICTAGACTTGCATAATAAGAATGTTAAGATGCTATGGCTAAA 1080	
QY	361 IleGlnAlaLysIleProGlyLeuLysArgLysAlaGlu 373	
Db	1081 ATCCAAGCAAAATCCCTGGATTGAAGCGCAAGCTGAA 1119	
RESULT 2		
HSA251384		
LOCUS	Homo sapiens mRNA for Nogo-B protein (Nogo gene).	linear PRI 22-JUL-2000
DEFINITION		
ACCESSION	AJ251384	
VERSION	AJ251384.1 GI:9408097	
KEYWORDS	Nogo gene; Nogo-B protein.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
REFERENCE	1	
AUTHORS	Prinjha, R., Moore, S.E., Vinson, M., Blake, S., Morrow, R.,	
TITLE	Christie, G., Michalovich, D., Simmons, D.L. and Walsh, F.S.	
JOURNAL	Inhibitor of neurite outgrowth in humans	
MEDLINE	Nature 403 (6768), 383-384 (2000)	
REFERENCE	2 (bases 1 to 1122)	
AUTHORS	Michalovich, D.	
TITLE	Direct Submission	
JOURNAL	Submitted (29-NOV-1999) Michalovich D., Bioinformatics, Smithkline	
	Beecham, Third Avenue, Harlow, Essex, CM19 5AW, BOSNIA AND	
	HERZEGOVINA	
FEATURES	Location/Qualifiers	
source	1..1122	
	/organism="Homo sapiens"	
	/mol_type="mRNA"	
	/db_xref="taxon:9606"	
gene	1..1122	
CDS	1..1122	
	/gene="Nogo"	
	/gene="Nogo"	
	/codon_start=1	
	/product="Nogo-B protein"	
	/protein_id="CAB93249.1"	
	/db_xref="GI:9408098"	
	/db_xref="GOA:Q9NQC3"	
	/db_xref="UniProt/Swiss-Prot:Q9NQC3"	
	/translation="MEDLDQSLVSSSDSPRPQPAFKYQVFRFEDDEEEDEEED	
	EEDELELEVLERKEPAAGLSAPVTPAAGAPLMDFGNDFVPPRGLPAAPVPAP	
	EROPSDPVSSTVAPSLASAAVSPSKLPEDDEPPARPPPPPPASVSPQAEFVWT	
	PPAPAAPSTPAAPKRRGGSGVVVDLLVWRDIKKTGVVFGASLFLLLSLTVFSIV	
	SVTAYIALALLSLVTSIFRIYKGVICAIQKSDGHPHPRAYLESEVALSEELVKYNSA	
	LGHVNCTIKELRLFLVDLVDLSLRFALVMVVFVTVGALFNLTLILLALLISLFSVPV	
	IYERHOAQIDHYLGLANKNVKDMAKIQAQIPGLKKRAE"	
ORIGIN		

## Alignment Scores:

Pred. No.: 2,97e-67 Length: 1122  
 Score: 1879.00 Matches: 370  
 Percent Similarity: 99.20% Conservatism: 0  
 Best Local Similarity: 99.20% Mismatches: 3  
 Query Match: 98.38% Indels: 0  
 DB: 9 Gaps: 0

US-09-544-776-2 (1-373) x HSA251384 (1-1122)

QY 1 MetGluAspLeuArgGlnSerProLeuValSerSerSerSerSerProProArgProGln 20  
 DB 1 ATGGAACACCTGGACCAAGTCTCTCTGCTCGTCTCGACAGCCACCCCGCGCCGAC 60  
 QY 21 ProAlaPheLysTyrGlnPheValArgGluProGluAspGluGluGluGluGlu 40  
 DB 61 CCCGCTTCAAGTACCAAGTCTGCTGAGGAGCCCGGAGGAGGAGGAGGAGGAG 120  
 QY 41 GluGluGluAspGluAspGluAspGluGluGluValLeuGluArgLysProAla 60  
 DB 121 GAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 180  
 QY 61 AlaGlyLeuSerAlaAlaProValProThrAlaProAlaAlaGlyAlaProLeuMetAsp 80  
 DB 181 GCCGGGCTGTCGCGGCCCGCCAGTCCACCGCCCTCGCGCGCGCCCTGATGGAC 240  
 QY 81 PheGlyAsnAspPheValProProAlaProArgGlyPheLeuProAlaAlaProProVal 100  
 DB 241 TTCGAAATGACTCTGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGTC 300  
 QY 101 AlaProGluArgGlnProSerTrpAspProSerProValSerSerThrValProAlaPro 120  
 DB 301 GCCCGGAGCGGAGCGGCTCTTGGGACCGGAGCGCGTGTCTGCGCGCGCGCGCA 360  
 QY 121 SerPheLeuSerAlaAlaAlaValSerProSerLysLeuProGluAspGluProPro 140  
 DB 361 TCCCGCGTGTCTGTCGCGCGAGTCTCGCCCTCCCAAGCTCCTGAGGAGCGAGCGCTCG 420  
 QY 141 AlaArgProProProProProAlaSerValSerProGlnAlaGluProValTrpThr 160  
 DB 421 GCCCGGCT 480  
 QY 161 ProProAlaProAlaProAlaProSerThrProAlaProLysArgArgGly 180  
 DB 481 CCGCAGCCCGGCT 540  
 QY 181 SerSerGlySerValValValValValValValValValValValValValValVal 200  
 DB 541 TCCCTCGGCTCAGTGGTGTGTGACCTCTGTGAGGAGACATTAAGAAGACTGGAGTG 600  
 QY 201 ValPheGlyAlaSerLeuPheLeuLeuSerLeuThrValPheSerIleValSerVal 220  
 DB 601 GTGTTGGTGGCCGCTATTCCTGCTCTTCAATGACAGTATTCAGCATTTGAGCGTGA 660  
 QY 221 ThrAlaTyrIleAlaLeuAlaLeuLeuSerValThrIleSerProArgIleTyrLysGly 240  
 DB 661 ACAGCTACATTCGCTTGGCCCTGCTCTGTGACCATCAGCTTAGGATATACAAGGT 720  
 QY 241 ValIleGlnAlaIleGlnLysSerAspGluGlyHisProPheArgAlaTyrLeuGluSer 260  
 DB 721 GTGATCCAGCTATCCAGAAATCAGATGAAGGCCACCATTCAGGGCATATCTGGAATCT 780  
 QY 261 GluValAlaIleSerGluGluLeuValGlnLysTyrSerAsnSerAlaLeuGlyHisVal 280  
 DB 781 GAAGTCTATATCTGAGGAGTGGTTGATCAGAGTACAGTAACTCTGCTTGTGTCATGTG 840  
 QY 281 AsnCysThrLysGluLeuArgArgLeuPheLeuValAspLeuValAspSerLeu 300  
 DB 841 AACTGCACGATAAAGAACTCAGCGGCTCTTCTTAGTGTATGATTTAGTTATCTCTG 900  
 QY 301 LysPheAlaValLeuMetTrpValPheThrTyrValGlyAlaLeuPheAsnGlyLeuThr 320

Db 901 AAGTTTCAGTGTGTGATGTGGGTATTTACCATGTGTGGTGCCTGTTTAAATGGTCTGACA 960  
 QY 321 LeuLeuIleLeuAlaLeuIleSerLeuPheSerValProValIleTyrGluArgHisGln 340  
 Db 961 CTACTGATTTGGCTCTCATTTCACTTCCTTCAGTGTTCCTGTTATTTATGACGGCATCAG 1020  
 QY 341 AlaGlnIleAspHisTyrLeuGlyLeuAlaAsnLysAsnValLysAspAlaMetAlaLys 360  
 Db 1021 GCACAGATAGATCATTTATCTAGGACTTGCATAAAGAATGTTAAAGATGCTATGGCTAAA 1080  
 QY 361 IleGlnAlaLysIleProGlyLeuLysArgLysAlaGlu 373  
 Db 1081 ATCCAAGCAAAATCCCTGGATTGAAGCGCAAGACTGAA 1119

## RESULT 3

LOCUS BC010737 1485 bp mRNA linear PRI 29-JUN-2004  
 DEFINITION Homo sapiens reticulon 4, transcript variant 2, mRNA (cDNA clone IMAGE:3901353), complete cds.

ACCESSION BC010737

VERSION BC010737

KEYWORDS BC010737.1 GI:14789600

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1485)

AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Dege, J.G.,

Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,

Altschul, S.F., Zebberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,

Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,

Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,

Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,

Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,

Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,

Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J.,

McKernan, R., Malek, J.A., Gunaratne, P.H., Richards, S.,

Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,

Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,

Fahney, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,

Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D.,

Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,

Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalusz, D.E.,

Schmerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length

human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932

2 (bases 1 to 1485)

Strausberg, R.

Direct Submission

Submitted (12-JUL-2001) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)

Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)

Center, Stanford University School of Medicine, Stanford, CA 94305

Web site: <http://www-ahgc.stanford.edu>

Contact: (Dickson, Mark) [mcd@paxil.stanford.edu](mailto:mcd@paxil.stanford.edu)

Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,

R. M.

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAK Plate: 14 Row: n Column: 2

This clone was selected for full length sequencing because it

passed the following selection criteria: Hexamer frequency ORF

analysis  
This clone has the following problem: The cds is short compared to  
the longest cds in the locus.

## FEATURES

## source

## Location/Qualifiers

```

1. .1485
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_image="3901353"
/issue_type="Pancreas, epithelioid carcinoma"
/clone_lib="NIH MGC 70"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1. .1485
/genes="RTN4"
/note="synonyms: NSP-CL, NOGO, ASY, NI220/250, NSP, RTN-X"
/db_xref="LocusID:57142"
/db_xref="MIM:604475"
6. .1127
/genes="RTN4"
/codon_start=1
/product="reticulon 4, isoform B"
/protein_id="AAH10737.1"
/db_xref="GI:14789601"
/db_xref="LocusID:57142"
/db_xref="MIM:604475"
/translation="MEDLDQSLVSSDSPPRPQPAFKYQVFRPEDEEEEBEED
EDLELELEKPAAGLAFVTPAAGAPLMDFGNDFVPPAPRGFLSPAAPFVAP
ERQPSWDSPVSTVPASPLSAAAVSPSKLPEDDEPPARPPPPASVSPAQBVMV
PPAPAPPSPTPAAPKRGSSGVVDLLYWRDKKTGVVFGASLFLLLSLTVFSIV
SVTAYTALALSVTISFRLYKGVIOAIOKSDSGHPFRAYLESEVALSEELVOKYSNA
LGHVNTIKELRRLFLVDLDSLPFVLMWFTYVGFALFNGLLTLLILALISLFSVPV
IYERHOQAIDHYGLGNKVNVDMAKIQKIPGLKRKAE"

```

## ORIGIN

```

Alignment Scores:
Pred. No.:      3 91e-67      Length:      1485
Score:          1879.00      Matches:      370
Percent Similarity: 99.20%      Conservative: 0
Best Local Similarity: 99.20%      Mismatches: 3
Query Match:      98.38%      Indels:      0
DB:               9          Gaps:      0

```

US-09-544-776-2 (1-373) x BC010737 (1-1485)

```

QY 1 MetGluAspLeuAspGlnSerProLeuValSerSerSerAspSerProProArgProGln 20
DB 6 ATGGAAGACCTGGACCACTCTCTCTGTCTCGTCTCGACAGCCACCCCGCGCGCAG 65
QY 21 ProAlaPheLysTyGlnPheValArgGluProGluAspGluGluGluGluGlu 40
DB 66 CCCGCGTTCAAGTACCAGTTCGTGAGGAGCCGAGGACGAGGAGGAAGAGGAGGAG 125
QY 41 GluGluGluAspGluAspGluAspGluGluGluGluGluGluGluGluGluGlu 60
DB 126 GAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 185
QY 61 AlaGlyLeuSerAlaAlaProValProThrAlaProAlaAlaGlyAlaProLeuMetAsp 80
DB 186 GCGCGGCTGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 245
QY 81 PheGlyAsnAspPheValProProAlaProArgGlyPheLeuProAlaAlaProVal 100
DB 246 TTCGGAATATGACTTCGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 305
QY 101 AlaProGluArgGlnProSerTrpAspProSerProValSerSerThrValProAlaPro 120
DB 306 GCGCGGAGCGGAGCGCGCTTGGGACCGAGCCGCGGTCTGTCGACCGTCCCGCGCA 365
QY 121 SerPheLeuSerAlaAlaValSerProSerLysLeuProGluAspAspGluProPro 140
DB 366 TCCCGCTGTCTGTGTCGCGCGCTTCCAGCTCCCTGAGGACGACGAGCGCTCGG 425

```

```

QY 141 AlaArgProProProProProAlaSerValSerProGlnAlaGluProValTrpThr 160
DB 426 GCGCGGCTTCCCGCTTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 485
QY 161 ProProAlaProAlaProAlaProProSerThrProAlaAlaProLysArgArgGly 180
DB 486 CCGCGAGCGCGCGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 545
QY 181 SerSerGlySerValValValAspLeuLeuTyTrpArgAspLeuLysLysThrGlyVal 200
DB 546 TCCTCGGCTCAGTGTGTGTTGACCTCTCTGACTCGAGAGACATTAAAGAAGACTGGAGTG 605
QY 201 ValPheGlyAlaSerLeuPheLeuLeuSerLeuThrValPheSerLeuValSerVal 220
DB 606 GTGTTTGTGTCAGCGCTATTCTCTGCTCTTCAATGACAGATTATCAGCAATTGTGACGTA 665
QY 221 ThrAlaTrpIleAlaLeuAlaLeuSerValThrIleSerProArgIleTrpLysGly 240
DB 666 ACAGCTACATGCTTGGCCCTGCTCTCTGACCATCAGCTTTAGGATATACAAAGGT 725
QY 241 ValIleGlnAlaIleGlnLysSerAspGluGlyHisProPheArgAlaTrpLeuGluSer 260
DB 726 GTGATCCAAGCTATCCAGAAATCAGATGAAGCCACCCATTCAAGGCATATCTGGAATCT 785
QY 261 GluValAlaIleSerGluLeuValGlnLysTyTrpSerAsnSerAlaLeuGlyHisVal 280
DB 786 GAAGTTGTCTATATCTGAGAGGTGTGTTCAAGATGACAGTAATCTGCTCTTGGTCATGTG 845
QY 281 AenCysThrIleLysGluLeuArgArgLeuPheLeuValAspLeuValAspSerLeu 300
DB 846 AACTGCACGATAAAGAACTCAGGCGCTCTCTCTTAGTTGATGATTAGTTAGTTCTCTG 905
QY 301 LysPheAlaValLeuMetTrpValPheThrTyValGlyAlaLeuPheAsnGlyLeuThr 320
DB 906 AAGTTTGCAGTGTTCATGTGGTATTTACTATGTGTGTCCTGTTTAAATGCTGTGACA 965
QY 321 LeuLeuIleAlaLeuIleSerLeuPheSerValProValIleTrpGluArgHisGln 340
DB 966 CTACTGATTTTGGCTCTCATTTCACTTCAGTGTCTCTGTTATTTATGAACGGCATCAG 1025
QY 341 AlaGlnIleAspHisTyLeuGlyLeuAlaAsnLysAsnValLysAspAlaMetAlaLys 360
DB 1026 GCACAGATAGATCATTTATAGGACTTGCATAATAAGATGTTAAAGATGCTATGCTAAA 1085
QY 361 IleGlnAlaLysIleProGlyLeuLysArgLysAlaGlu 373
DB 1086 ATCCAAGCAAAAATCCCTGGATTGAAGCGCAAGCTGAA 1124
RESULT 4
LOCUS BD231889 1610 bp DNA linear PAT 17-JUL-2003
DEFINITION Bone marrow secreted proteins and polynucleotides.
ACCESSION BD231889
VERSION BD231889.1 GI:33041659
KEYWORDS JP 2002511231-A/3.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1610)
AUTHORS Lin, H. and Cao, L.
TITLE Bone marrow secreted proteins and polynucleotides
JOURNAL Patent: JP 2002511231-A 3 16-APR-2002;
COMMENT CHIRON CORP
OS Homo sapiens (human)
PN JP 2002511231-A/3
PD 16-APR-2002
PF 18-DEC-1998 JP 2000526635
PR 30-DEC-1997 US 60/068958, 24-SEP-1998 US 60/101603 PR
30-SEP-1998 US 60/102540
PI HAISHAN LIN, LI CAO
PC C12N15/09, A61K38/00, A61P43/00, C07K14/47, C07K16/18, C12N5/10, PC

```

C12Q1/68,  
PC C12N15/00.A61K37/02.C12N5/00  
CC Bone marrow secreted proteins and polynucleotides FH Key  
Location/Qualifiers  
FT source 1..1610  
FT Location/Qualifiers  
1..1610  
/organism="Homo sapiens (human)"  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

FEATURES  
source  
Alignment Scores:  
Pred. No.: 4,22e-67 Length: 1610  
Score: 1879.00 Matches: 370  
Percent Similarity: 99.20% Conservativeness: 0  
Best Local Similarity: 99.20% Mismatches: 3  
Query Match: 98.38% Indels: 0  
DB: Gaps: 0

US-09-544-776-2 (1-373) x BD231889 (1-1610)

QY 1 MetGluAspLeuAspGlnSerProLeuValSerSerSerAspSerProProArgProGln 20  
DB 132 ATGGAAGACCTGGACCACTCTCTGCTCGTCTCGGACAGCCACCCCGCGCGCAG 191  
QY 21 ProAlaPheIleValArgGluProGluAspGluGluGluGluGluGlu 40  
DB 192 CCCGCGTTCAAGTACCAGTTCTGTGAGGAGCCCGAGGAGGAGGAGGAGGAG 251  
QY 41 GluGluGluAspGluAspGluAspGluGluGluGluGluGluGluGluGlu 60  
DB 252 GAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 311  
QY 61 AlaGlyLeuSerAlaAlaProValProThrAlaProAlaAlaGlyAlaProLeuMetAsp 80  
DB 312 GCCGGCGTGTCCGCGCGCCAGTGGCCAGCCCGCCCGCGCGCGCGCGCGCG 371  
QY 81 PheGlyAsnAspPheValProAlaProArgGlyPheLeuProAlaProVal 100  
DB 372 TTCGGAATGACTGTGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 431  
QY 101 AlaProGluArgGlnProSerTrpAspProSerProValSerSerThrValProAlaPro 120  
DB 432 GCCCGGAGCGGAGCGGCTCTGGGACCGGAGCGCGGTGTCGACCGCGCGCCA 491  
QY 121 SerPheLeuSerAlaAlaValSerProSerLysLeuProGluAspAspGluProPro 140  
DB 492 TCCCGCGTGTCTGTCGCGGAGTCTCGCCCTCCAAGCTCCCTGAGGACGACGCGCTCCG 551  
QY 141 AlaArgProProProProProAlaSerValSerProGlnAlaGluProValTrpThr 160  
DB 552 GCCCGGCGTCCCTCTCCCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 611  
QY 161 ProProAlaProAlaProAlaProProSerThrProAlaAlaProLysArgArgGly 180  
DB 612 CCGCCAGCGCGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 671  
QY 181 SerSerGlySerValValValAspLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 200  
DB 672 TCCTCGGGCTCAGTGGTGTGTACCTCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAG 731  
QY 201 ValPheGlyAlaSerLeuPheLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 220  
DB 732 GTGTTGGTCCAGCCATTCTCTGCTCTTTCATGACATATACAGATTGAGCGGTA 791  
QY 221 ThrAlaTrpIleAlaLeuAlaLeuLeuSerValThrIleSerProArgIleTrpLysGly 240  
DB 792 ACAGCCTACATTGCTTGGCCCTGCTCTCTGTGACCATCAGCTTAGGATATACAAGG 851  
QY 241 ValIleGlnAlaIleGlnLysSerAspGluGlyIleProPheArgAlaTrpLeuGluSer 260

DB 852 GTGATCCAAAGCTATCCAGAAATCAGATGAAGGCCACCCATTCAGGCGATATCTGGAACT 911  
QY 261 GluValAlaIleSerGluGluLeuValGlnLysTrpSerAsnSerAlaLeuGlyHisVal 280  
DB 912 GAAGTTGCTATATCTGAGGAGTGGTTTCAGAGTACAGTAATCTCTCTTGGTCATGTC 971  
QY 281 AsnCysThrIleLysGluLeuArgArgLeuPheLeuValAspAspLeuValAspSerLeu 300  
DB 972 AACTGCACGATAAAGCACTCAGGCGCCTCTCTTAGTTAGTATGATTTAGTTCTCTG 1031  
QY 301 LysPheAlaValLeuMetTrpValPheThrTrpValGlyValAlaLeuPheGlnGlyLeuThr 320  
DB 1032 AAGTTTGCAGTGTGATGTGGGTATTTACTATGTTGGTGGCTTGTATATGCTCTGACA 1091  
QY 321 LeuLeuIleLeuAlaLeuIleSerLeuPheSerValProValIleTrpGluArgHisGln 340  
DB 1092 CTACTGATTTGGCTCTCATTTTCACCTCTTCAGTGTCTCTGTTATTTATGAACGCGATCAG 1151  
QY 341 AlaGlnIleAspHisTrpLeuGlyLeuAlaAsnLysAsnValLysAspAlaMetAlaLys 360  
DB 1152 GCACAGATAGATCAATTATCTAGGACTTGCAATTAAGATGTTAAAGATGCTATGGCTAAA 1211  
QY 361 IleGlnAlaIleValProGlyLeuLysArgLysAlaGlu 373  
DB 1212 ATCCAAGCAAAATCCCTGGATTGAAGCGCAAGCTGAA 1250

RESULT 5  
AK129806 1619 bp mRNA linear PRI 10-SEP-2003  
LOCUS Homo sapiens cDNA FLJ26295 fis, clone DMC07157.  
DEFINITION AK129806  
ACCESSION AK129806.1 GI:34526422  
VERSION oligo capping; fis (full insert sequence).  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 Ota, T., Nakagawa, S., Senoh, A., Mizuguchi, H., Inagaki, H., Suzuki, Y., Hata, H., Nakagawa, K., Mizuno, S., Morinaga, M., Kawamura, M., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Nishikawa, T., Sugiyama, A., Kawakami, B., Nagai, K., Isogai, T. and Sugano, S. NEDO human cDNA sequencing project  
Unpublished  
REFERENCE 2 (bases 1 to 1619)  
Sugano, S. and Suzuki, Y.  
Direct Submission  
Submitted (31-JUL-2003) Sumio Sugano, Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure, Human Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan (E-mail: fldcna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286, Fax:81-3-5449-5416)  
COMMENT NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction and 5'-end one pass sequencing: Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure, Human Genome Center; 3'-end one pass sequencing: RAB; clone selection for full insert sequencing: RAB and Helix Research Institute.  
FEATURES  
source  
1..1619  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="DMC07157"  
/tissue\_type="dermoid tumor"  
/clone\_lib="DMC"  
/note="cloning vector: pME18SFL3"

ORIGIN  
Alignment Scores:  
Pred. No.: 4,25e-67 Length: 1619  
Score: 1879.00 Matches:



Percent Similarity: 99.20% Conservative: 0  
 Best Local Similarity: 99.20% Mismatches: 3  
 Query Match: 98.38% Indels: 0  
 DB: 9 Gaps: 0

US-09-544-776-2 (1-373) x AK129806 (1-1619)

QY	1	MetGluAspLeuAspClnSerProLeuValSerSerSerSerSerSerProProArgProGln	20
DB	158	ATGGAAGACCTGGACCAAGCTCTCTCTGGTCTCGTCTCGACAGCCACCCCGCGCGAG	217
QY	21	ProAlaPheLysTyrglnPheValArgGluProGluAspGluGluGluGluGluGlu	40
DB	218	CCGCGTTCAAGTACCAAGTTCGTGAGGAGCCGAGGACGAGGAGGAGGAGGAG	277
QY	41	GluGluGluAspGluAspGluAspLeuGluGluGluGluGluGluGluGluGluGlu	60
DB	278	GAAGAGGAGGACGAGGACCAAGACCTGGAGGAGCTGGAGGTCTGGAGAGGAGCCCGC	337
QY	61	AlaGlyLeuSerAlaAlaAlaProValProThrAlaProAlaAlaGlyAlaProLeuMetAsp	80
DB	338	GCCGGGCTCTCGCGGCGCCAGTGCACCCGCTCCGCGCGGCGCGCCCTGATGGAC	397
QY	81	PheGlyAsnAspPheValProProAlaProArgGlyPheLeuProAlaAlaProProVal	100
DB	398	TTCCGNAATGATTCGTGCGCGCGCGCGCGCGCGCGGACCCCTCGCGCGCGCTCCCGCCGTC	457
QY	101	AlaProGluArgGlnProSerTrpAspProSerProValSerSerThrValProAlaPro	120
DB	458	GCCCGGAGCGGACGCGCTTTGGGACCGAGCCGCGTCTGCTGACCGCGCGCGCA	517
QY	121	SerPheLeuSerAlaAlaAlaValSerProSerLysLeuProGluAspGluProPro	140
DB	518	TCCCGGCTCTGCTGCGCGAGTCTGCGCCTCCAACTCCCTGAGGACACGAGCGCTCGG	577
QY	141	AlaArgProProProProProAlaSerValSerProGlnAlaGluProValTrpThr	160
DB	578	GCCCGGCTCCCGCTCCCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	637
QY	161	ProProAlaProAlaProAlaProProSerThrProAlaAlaProLysArgArgGly	180
DB	638	CGCGCAGCCCGCGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	697
QY	181	SerSerGlySerValValValAspLeuLeuTyrrTrpArgAspIleIleIleIleIleIleIle	200
DB	698	TCCTCGGCTCAGTGTGTGACCTCTGTTACTGAGAGACATTAAAGAACTCGAGTG	757
QY	201	ValPheGlyAlaSerLeuPheLeuLeuLeuSerLeuThrValPheSerIleValSerVal	220
DB	758	GTGTTTGGTGGCAGCCTATTCTGCTGCTTTTCATTGACAGTATTGAGCATTGAGCGTA	817
QY	221	ThrAlaTyrlleAlaLeuAlaLeuLeuSerValThrIleSerProArgIleTyrlsGly	240
DB	818	ACAGCGCTACATTGCTTGGCCCTGCTCTGTGACCATCAGCTTTAGGATATACAGGGT	877
QY	241	ValIleGlnAlaIleGlnLysSerAspGluGlyHisProPheArgAlaTyrlleuGluSer	260
DB	878	GTGATCCAGCTATCCAGAAATCAGATGAAGCCACCCATTGAGGACATTCGGAATCT	937
QY	261	GluValAlaIleSerGluGluLeuValGlnLysTyrrSerAsnSerAlaLeuGlyHisVal	280
DB	938	GAAGTTGTATATCTCAGGAGTGGTTTCAGAAAGTACAGTAATCTGCTTGTCTCATGTG	997
QY	281	AnCysThrIleLysGluLeuArgArgLeuPheLeuValAspLeuValAspSerLeu	300
DB	998	AACTGCAGATAAAGAACTCAGGGCGCTCTTTCTAGTTGATGATTTAGTTGATTTCTCTG	1057
QY	301	LysPheAlaValLeuMetTrpValPheThrTyrrValGlyAlaLeuPheAsnGlyLeuThr	320
DB	1058	AAATTTCAGTGTGTGATGAGGATTTACCTATGTTGTCCTTGTGTTATGCTGTGACA	1117
QY	321	LeuLeuIleLeuAlaLeuIleSerLeuPheSerValProValIleTyrrGluArgHisGln	340

DB	1118	CTACTGATTTGGCTCTCATTTCACTCTTCAGTGTTCCTGTTATTATTATGAACGGCATCAG	1177
QY	341	AlaGlnIleAspHisTyrlleuGlyLeuAlaAsnLysAsnValLysAspAlaMetAlaLys	360
DB	1178	GCACAGATAGATCATTTCTAGGACTTGCAATAAAGAAATGTTAAAGATGCTATGGCTAAA	1237
QY	361	IleGlnAlaLysIleProGlyLeuLysArgLysAlaGlu	373
DB	1238	ATCCAAGCAAAAATCCCTGGATTGAAGCGCAAGCTGAA	1276

# RESULT 6

BC012619	1654 bp	mRNA	linear	PRI 29-JUN-2004
Homo sapiens				
IMAGE:40827561				
BC012619				
BC012619.1	GI:15214977			
Homo sapiens				
Homo sapiens				
ORGANISM				
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
1 (bases 1 to 1654)				
REFERENCE				
AUTHORS				

Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hong, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaby, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywinski, M.I., Skalska, U., Smalish, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.  
 Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences  
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

JOURNAL	1247932
PUBMED	2 (bases 1 to 1654)
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	

Strausberg, R.  
 Direct Submission  
 Submitted (15-AUG-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA  
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 Contact: MGC help desk  
 Email: [cgapbs-re@mail.nih.gov](mailto:cgapbs-re@mail.nih.gov)  
 Tissue Procurement: ATCC  
 CDNA Library Preparation: CLONTECH Laboratories, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Institute for Systems Biology  
<http://www.systemsbio.org>  
 contact: [amadnan@systemsbio.org](mailto:amadnan@systemsbio.org)  
 Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Series: IRAL Plate: 19 Row: d Column: 8  
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 24431932  
 This clone has the following problem: The cds is short compared to the longest cds in the locus.  
 Location/Qualifiers  
 1. 1654

FEATURES  
 source



```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4082756"
/tissue_type="Brain, glioblastoma"
/clone_lib="NIH_MGC_57"
/lab_host="DH10B"
/notes="Vector: pDNR-LIB"
1. .1654
/genes="RTN4"
/notes="synonyms: NSP-CL, NOGO, ASY, NI220/250, NSP, RTN-X"
/db_xref="LocusID:57142"
/db_xref="MIM:604475"
156. .1277
/genes="RTN4"
/codon_start=1
/product="reticulon 4, isoform B"
/protein_id="AAH12619.1"
/db_xref="GI:15214978"
/db_xref="LocusID:57142"
/db_xref="MIM:604475"
/translators="MEDDQSLVSSSDSPRPQPAKYQFVREPEDEEBEERED
EDDLRELEVLKPKAGLSAAPPVAPAGALPMDGNDPVPAPRGLPAPAPPAP
ERQSPNDPVSSTVAPSLSAANVPSKLPEDDEPPAPPPPPASVSPQEPWT
PPAPAPAPSTTAAKRRSGSVVDLLIWRDIKKTGVVFGASLFLLSLTVFSIV
SVTAYTALALSLVTIIFRIYKGVIAIQKSDGHPFRAVLESEVALSEELVKYSNSA
LGHVNCITKELRLFLVDLVDLSKFAVLMMVFTYVVGALFNGLLILALISLFSVPV
IYERHQAIIDHYLGANKVNDAMAKIQAKIPLKPKAE"

ORIGIN
Alignment Scores:
Pred. No.: 4.34e-67 Length: 1654
Score: 1879.00 Matches: 370
Percent Similarity: 99.20% Conservative: 0
Best Local Similarity: 99.20% Mismatches: 3
Query Match: 98.38% Indels: 0
DB: 9 Gaps: 0

US-09-544-776-2 (1-373) x BC012619 (1-1654)
QY 1 MetGluAspLeuAspGlnSerProLeuValSerSerSerAspSerProProArgProGln 20
DB 156 ATGGAAGACCTGGACCAAGTCTCTCTGCTCTGCTCTCGACAGCCACCCCGCGCGCAG 215
QY 21 ProAlaPheLysTyrGlnPheValArgGluProGluAspGluGluGluGluGlu 40
DB 216 CCGCGGTTCAAGTACCAAGTTCGTGGAGGAGCCCGAGGACGAGGAGGAGGAGGAG 275
QY 41 GluGluGluAspGluAspGluAspGluGluGluGluGluGluGluGluGluGlu 60
DB 276 GAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 335
QY 61 AlaGlyLeuSerAlaAlaProValProThrAlaProAlaAlaGlyAlaProLeuMetAsp 80
DB 336 GCCCGGGTGTCCGGCGGCCCGAGTGCACCCCGCTGCGCGCGCGCGCGCGCGCGGAC 395
QY 81 PheGlyAsnAspPheValProProAlaProArgGlyPheLeuProAlaAlaProProVal 100
DB 396 TTCGGAATGACTTCGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGTCC 455
QY 101 AlaProGluArgGlnProSerTrpAspProSerProValSerSerThrValProAlaPro 120
DB 456 GCCCGGAGGCGGAGCGGCGTCTTGGACCCGAGCGCGGTGCGTGGCGCGCGCGGCA 515
QY 121 SerPheLeuSerAlaAlaAlaValSerProSerLysLeuProGluAspGluProPro 140
DB 516 TCCCGCGTGTCTGCGCGCGAGTCTGCGCTTCAAGCTCCCTGAGGACGACGAGCTCCG 575
QY 141 AlaArgProProProProProProAlaSerValSerProGlnAlaGluProValTrpThr 160
DB 576 GCCCGGCGTCTCTCTCCCGCGCGAGGAGTGGAGCCCCGAGGAGCGCGTGTGGAC 635
QY 161 ProProAlaProAlaProAlaAlaProProSerThrProAlaAlaProLysArgGly 180

```

```

636 CGCCAGCCCGGCTCCCGCGCGCGCCCTCCACCCCGCGCGCGCGCGCGCGCGCGC 695
181 SerSerGlySerValValValValValValValValValValValValValValVal 200
696 TCCTCCGGCTCAGTGGTGTGTGACCTCCTGTTACTGGAGAGACATTAGAGAGCTGGAGTG 755
201 ValPheGlyAlaSerLeuPheLeuLeuLeuSerLeuThrValPheSerLeuValSerVal 220
756 GTGTTTGGTGCCAGCCTATTCTCTGCTGCTTTTCATTGACAGTATTCAGCATTGTGAGCGTA 815
221 ThrAlaTyrIleAlaLeuAlaLeuLeuSerValThrIleSerProArgIleTyrLysGly 240
816 ACAGCCTACATTGCTTGGCCCTGCTCTCTGTGACCATCAGCTTTAGGATATACAAGGCT 875
241 ValIleGlnAlaIleGlnLysSerAspGluGlyHisProPheArgAlaTyrLeuGluSer 260
876 GTGATCCAGCTATCCAGAAATCAGATGAAGCCACCATTCAGGCGCATATCTGGAACTCT 935
261 GluValAlaIleSerGluGluLeuValGlnLysTyrSerAsnSerAlaLeuGlyHisVal 280
936 GAAGTTGCTATATCTCAGGAGTGGTTTCAAGAGTACAGTAATTTCTGCTTGGTCAATGTG 995
281 AsnCysThrIleLysGluLeuArgArgLeuPheLeuValAspAspLeuValAspSerLeu 300
996 AACTGCACCATTAAGGAACCTCAGGCGCCCTCTCTTAGTTGATGATTAGTTGATTCTCTG 1055
301 LysPheAlaValLeuMetTrpValPheThrTyrValGlyValAlaLeuPheAsnGlyLeuThr 320
1056 AGTTTGGCAGTGTGATGTTGGGTATTTACTATGTTGGTGGCTGTTTAAATGTTCTGACA 1115
321 LeuLeuIleLeuAlaLeuIleSerLeuPheSerValProValIleTyrGluArgHisGln 340
1116 CTACTGATTTGGCTCTCATTTTCACTCTTCAGTGTTCCTGTTTATTATGAACGGCATCAG 1175
341 AlaGlnIleAspHisTyrLeuGlyLeuAlaAsnLysAsnValLysAspAlaMetAlaLys 360
1176 GCACAGATAGATCATTTATCTAGGACTTGCAAAATGAAGATGTTAAAGATGCTATGGCTAA 1235
361 IleGlnAlaLysIleProGlyLeuLysArgLysAlaGlu 373
1236 ATCCAAGCAAAATCCCTGATGGAAGCGCAAGCTGAA 1274

RESULT 7
LOCUS BC026788 1668 bp mRNA linear PRI 29-JUN-2004
DEFINITION Homo sapiens reticulon 4, transcript variant 2, mRNA (cdna clone
IMAGE:3862911), complete cds.
ACCESSION BC026788
VERSION BC026788.1 GI:20070661
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1668)
AUTHORS
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altshul,S.P., Zeeberg,B., Buetow,K.H., Schaefer,C.P., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,T.E., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Ustin,T.B., Toshlyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettner,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,

```

Schmerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.  
 Generation and initial analysis of more than 15,000 full-length  
 human and mouse cDNA sequences  
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
 12477932  
 2 (bases 1 to 1668)  
 Strausberg, R.  
 Direct Submission  
 Submitted (04-APR-2002) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA  
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 Contact: MGC help desk  
 Email: [cgabbs-r@mail.nih.gov](mailto:cgabbs-r@mail.nih.gov)  
 Tissue Procurement: DCTD/DTF  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Baylor College of Medicine Human Genome  
 Sequencing Center  
 Center code: BCM-HGSC  
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>  
 Contact: [amg@bcm.tmc.edu](mailto:amg@bcm.tmc.edu)  
 Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H.,  
 Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,  
 A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Series: IRAK Plate: 21 Row: h Column: 13  
 This clone was selected for full length sequencing because it  
 passed the following selection criteria: Hexamer frequency ORF  
 analysis, similarity but not identity to protein  
 This clone has the following problem: The cds is short compared to  
 the longest cds in the locus.

FEATURES  
 source  
 1..1668  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:3862911"  
 /issue\_type="Ovary, adenocarcinoma"  
 /clone\_lib="NIH MGC 66"  
 /lab\_host="DH10B"  
 /not5="Vector: PCMV-SPORT6"  
 1..1668  
 /gene="RTN4"  
 /note="synonyms: NSP-CL, NOGO, ASY, NI220/250, NSP, RTN-X"  
 /db\_xref="LocusID:57142"  
 /db\_xref="MIM:604475"  
 187..1308  
 /gene="RTN4"  
 /codon\_start=1  
 /product="reticulon 4, isoform B"  
 /protein\_id="AAH26788.1"  
 /db\_xref="GI:20070662"  
 /db\_xref="LocusID:57142"  
 /db\_xref="MIM:604475"  
 /translation="MEDLDQSLVSSDPPRPQAFKQYQVREDEDEEBEED  
 EDELEELERKPAAGLSAAPTAPAAAGAPLMDFGNDPVPAPRGPLPAAPFPAP  
 EQPSWDSPSVSTVPAPSLSAASVSKLPEDDEPPARPPPPASVSPQAEVPT  
 PAPAPAPPSPTPAAPKRRGSGSVVDLLYWRDIKKTGVVFGASLFLLLSLTPFSIV  
 SVTAYALALSVTISFRLYKGVIOAIOKSDEGHPPRAYLESEVALISELVQKYSNSA  
 LGHVNCTIKELRRLEFVDDLSLKFAPVLMVWFTVVGALFNLGLTLLILALISLFSVPV  
 IYERHQAOIDHYGLANKNVKDMAKIOAKIPGLKRKAE"

ORIGIN  
 Alignment Scores:  
 Pred. No.: 4,37e-67 Length: 1668  
 Score: 1879.00 Matches: 370  
 Percent Similarity: 99.20% Conservative: 0  
 Best Local Similarity: 99.20% Mismatches: 3  
 Query Match: 98.38% Indels: 0

DB: 9 Gaps: 0  
 US-09-544-776-2 (1-373) x BC026788 (1-1668)  
 QY 1 MetGluAspLeuAspGlnSerProLeuValSerSerSerSerProProArgProGln 20  
 DB 187 ATGGAAGACCTGGACCAAGTCTCTCTGGTCTCGTCTCGACAGCCACCCCGCGCGAG 246  
 QY 21 ProAlaPheLysTyrGlnPheValArgGluProGluAspGluGluGluGluGlu 40  
 DB 247 CCCCGCTTCAAGTACCACTTCGTGAGGAGCCCGGAGGACGAGGAGGAAGAGGAGAG 306  
 QY 41 GluGluGluAspGluAspGluAspLeuGluGluLeuValLeuGluArgLysProAla 60  
 DB 307 GAAGAGGAGGACGAGGACGAGACCTCGAGGAGCTGGAGGTCTGGAGAGGAGCCCGCC 366  
 QY 61 AlaGlyLeuSerAlaAlaProValProThrAlaProAlaAlaGlyAlaProLeuMetAsp 80  
 DB 367 GCCGGGCTGTCGCGGCCCCAGTGCACCGCCCTCGCGCGCGCGCCCTGATGGAC 426  
 QY 81 PheGlyAsnAspPheValProAlaProAlaProArgGlyPheLeuProAlaAlaProVal 100  
 DB 427 TTCGAAATGACTTCGTGCGCGCGCGCGCGCGCGCGCGCGCTGCCCGCGCGCTCCCCCGTC 486  
 QY 101 AlaProGluArgGlnProSerTyrAspProSerProValSerSerThrValProAlaPro 120  
 DB 487 GCCCGGAGCGGACGACCGCTCTTGGACCGGACCGCGCTGTGTGACCGTGCAGCGTGC 546  
 QY 121 SerPheLeuSerAlaAlaValSerProSerLysLeuProGluAspAspGluProPro 140  
 DB 547 TCCCGCTGTCTGTGCGCGAGTCTCGCCCTCCAAGCTCCCTGAGGACGACGAGGCTCGG 606  
 QY 141 AlaArgProProProProProAlaSerValSerProGlnAlaGluProValThr 160  
 DB 607 GCCCGGCTCCCGCTCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 666  
 QY 161 ProProAlaProAlaProAlaProProSerThrProAlaAlaProLysArgArgGly 180  
 DB 667 CGCCGAGCCCGGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 726  
 QY 181 SerSerGlySerValValValAspLeuLeuTyrTyrArgAspIleLysLysThrGlyVal 200  
 DB 727 TCCTCGGCTCAGTGGTGTGTGACCTCTCTGTACTGGAGAGACATTAAGAAGACTGGAGTG 786  
 QY 201 ValPheGlyAlaSerPheLeuLeuLeuSerLeuThrValPheSerIleValSerVal 220  
 DB 787 GTGTTGTGTGCAGCTATTCTCTGCTTTCAITGACAGTATTCAGCATTTGTGAGCGTA 846  
 QY 221 ThrAlaTyrIleAlaLeuAlaLeuLeuSerValThrIleSerProArgIleTyrLysGly 240  
 DB 847 ACAGCTTACATTCCTTGGCCCTGCTCTGTGTGACCATCAGCTTTAGGATATACAAAGGT 906  
 QY 241 ValIleGlnAlaIleGlnLysSerAspGluGlyHisProPheArgAlaTyrLeuGluSer 260  
 DB 907 GTGATCCAAGCTATCCAGAAATCAGATGAAGGCCACCCATTCAAGGCATATCTGGAATCT 966  
 QY 261 GluValAlaIleSerGluGluLeuValGlnLysTyrSerAsnSerAlaLeuGlyHisVal 280  
 DB 967 GAAGTGTCTATATCTGAGGAGTGTGTTCAGAGTACAGTAATCTGCTCTTGGTCATGTG 1026  
 QY 281 AsnCysThrIleLysGluLeuArgArgLeuPheLeuValAspAspLeuValAspSerLeu 300  
 DB 1027 AACTGCACGATAAAGAACTCAGGCGCTCTCTTCTTAGTTGATGATTAGTTAGTTCTCTG 1086  
 QY 301 LysPheAlaValLeuMetTyrValPheThrTyrValGlyAlaLeuPheAsnGlyLeuThr 320  
 DB 1087 AAGTTTGCAGTGTTCATGTGGGTATTATTACCTATGTGTGGTCCCTTGTAAATGCTGACA 1146  
 QY 321 LeuLeuIleLeuAlaIleSerLeuPheSerValProValIleTyrGluArgHisGln 340  
 DB 1147 CTACTGATTTTGGCTCTCATTTCACTTTCAGTGTCTCTGTTATTATTATGAACGCGATCAG 1206  
 QY 341 AlaGlnIleAspHisTyrLeuGlyLeuAlaAsnLysAsnValLysAspAlaMetAlaLys 360



```

LOCUS       BC068991               1728 bp      mRNA      linear      PRI 30-JUN-2004
DEFINITION  Homo sapiens reticulon 4, transcript variant 2, mRNA (cDNA clone
IMAGE:4634289), complete cds.
ACCESSION   BC068991
VERSION     BC068991.1  GI:46362518
KEYWORDS    '
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1728)
AUTHORS     Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zuber, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Ustin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzyzanski, M.I., Skalska, U., Smalusz, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
JOURNAL     Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
PUBMED      12477932
REFERENCE   2 (bases 1 to 1728)
AUTHORS     Strausberg, R.
TITLES      Direct Submission
JOURNAL     Submitted (05-APR-2004) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK      NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
Email: cgabs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca
Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield,
Andy Chan, Steve S. Chand, William Chow, Alison Cloutier, Ruth
Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao,
Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin,
Teika Olson, Diana Palmquist, Anca Petrescu, Anna Liisa Prabh,
Parvaneh Saeedi, JR Santos, Angélique Schnerch, Ursula Skalska,
Duane Smalusz, Jeff Stott, Miranda Tsai, George Yang, Jacquie
Schein, Asim Siddiqui, Rob Holt, Marco Marra.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov
Series: IRAL Plate: 54 Row: c Column: 22
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 28557782
This clone has the following problem: The cds is short compared to
the longest cds in the locus.
Location/Qualifiers
1. .1728
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4634289"
/tissue_type="Eye, retinoblastoma"
/clone_lib="NIH_MGC_16"

```

```

/lab_host="DH10B-P"
/note="Vector: pOTB7"
1. .1728
/gene="RTN4"
/note="synonyms: NSP-CL, NOGO, ASY, NI220/250, NSP, RTN-X"
/db_xref="LocusID:57142"
/db_xref="MIM:604475"
249. .1370
/gene="RTN4"
/codon_start=1
/product="reticulon 4, isoform B"
/protein_id="AAH68991.1"
/db_xref="GI:46362519"
/db_xref="LocusID:57142"
/db_xref="MIM:604475"
/translation="MED12QSPVSSSDSPRPQPAFKYQFVREPEDEEEED
EDEDLELEVERKEPAAGLSAAPTAAAGAPLMDFGNDPVPAPRGPLPAPVAP
EPQSDPSPVSTVAPSPLSAAVSPKLPEDDEPPAPPPPPSPVSPQAPVWT
PPAPAPAPSTPAAPKRGSGSVVDLLVWRDIKKTGVFGASLFLLSLTVSIV
SVTAYIALALLSIVTSIFRIYKGVIOAIQKSDGHPFRAYLEVAISELVQYNSA
LGHVNCITKELRRLFLVDLDFALFVLMVFTYVGFALFGLTLILLALLSLFVSPV
IYERHOAQIDHYLGLANKNVKAMAKIOAKIPGLKRKAE"
ORIGIN
Alignment Scores:
Pred. No.: 4,52e-67 Length: 1728
Score: 1879.00 Matches: 370
Percent Similarity: 99.20% Conservative: 0
Best Local Similarity: 99.20% Mismatches: 3
Query Match: 98.38% Indels: 0
DB: 9 Gaps: 0
US-09-544-776-2 (1-373) x BC068991 (1-1728)
QY 1 MetGluAspLeuAspGlnSerProLeuValSerSerSerProProArgProGln 20
DB 249 ATGGAAGACCTCGACCAAGTCTCTCTGGTCTCGTCTCGACAGCCACCCCGCGCAG 308
QY 21 ProAlaPheLeuTyGlnPheValArgGluProGluAspGluGluGluGluGlu 40
DB 309 CCCGCGTTCAAGTACCACTTCGTGAGGAGCCCGAGGAGGAGGAGGAGGAG 368
QY 41 GluGluGluAspGluAspGluLeuGluGluLeuGluValLeuGluArgLysProAla 60
DB 369 GAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 428
QY 61 AlaGlyLeuSerAlaAlaProValProThrAlaProAlaAlaGlyAlaProLeuMetAsp 80
DB 429 GCGCGGCTGTCGCGCGCCAGTGCACCGCCCTGCGCGCGCGCGCGCGCGCGCC 488
QY 81 PheGlyAsnAspPheValProAlaProAlaProAlaProAlaProAlaProVal 100
DB 489 TTCGGAATAGCTTCGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 548
QY 101 AlaProGluArgGlnProSerTrpAspProSerProValSerSerThrValProAlaPro 120
DB 549 GCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 608
QY 121 SerPheLeuSerAlaAlaAlaValSerProSerLysLeuProGluAspAspGluProPro 140
DB 609 TCCCGCTGTCTGTGTCGCGCAGTCTCGCCCTCAAGCTCCCTCGAGGAGGAGGAGGAG 668
QY 141 AlaArgProProProProAlaSerValSerProGlnAlaGluProValTrpThr 160
DB 669 GCGCGGCTCCCGCTCTCTCCCGCGCGAGCGTGAGCGCGCGCGCGCGCGCGCG 728
QY 161 ProProAlaProAlaProAlaProProSerThrProAlaAlaProLysArgGly 180
DB 729 CCGCGAGCGCGCGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 788
QY 181 SerSerGlySerValValValAspLeuLeuTyTrpArgAspIleIleIleValThrGlyVal 200
DB 789 TCTCGGGCTCAGTGGTGTGTTGACCTCTGTACTCGAGAGACATTAAGAAGACTCGAGTG 848

```

```
Qy 201 ValPheGlyAlaSerLeuPheLeuLeuLeuSerLeuThrValPheSerIleValSerVal 220
Db 849 GTGTTTGTCGCCAGCTTCTCTGCTCTTTCATTGACAGTATTACAGTATTGTCAGGGTA 908
Qy 221 ThrAlaTyrIleAlaLeuAlaLeuLeuSerValThrIleSerProArgIleTyrLysGly 240
Db 909 ACAGCCTACATTCGCTTGGCCCTGCTCTCTGTGACCATCAGCTTTAGGATATACAAGGT 968
Qy 241 ValIleGlnAlaIleGlnLysSerAspGluGlyHisProPheArgAlaTyrLeuGluSer 260
Db 969 GTGATCCAGCTATCCAGAAATCAGATGAAGCCACCANTTCAGGGCATATCTGGAAATCT 1028
Qy 261 GluValAlaIleSerGluGluLeuValGlnLysTyrSerAsnSerAlaLeuGlyHisVal 280
Db 1029 GAAGTTCCTATATCTGAGGAGTGTGTTTTCAGAGTACAGTAATCTGCTCTTGGTCATGTG 1088
Qy 281 AsnCysThrIleLysGluLeuArgArgLeuPheLeuValAspAspLeuValAspSerLeu 300
Db 1089 AACTGCACGATAAAGGAACCTCAGCGCCTCTTCTTAGTTGATGATTTAGTTGATCTCTG 1148
Qy 301 LysPheAlaValLeuMetTyrValPheThrTyrValGlyAlaLeuPheAsnGlyLeuThr 320
Db 1149 AAGTTTCAGTGTGATGTGGTATTTTACCTATGTTGGTGCCTTGTTTAATGGTCTGACA 1208
Qy 321 LeuLeuIleLeuAlaLeuIleSerLeuPheSerValProValIleTyrGluArgHisGln 340
Db 1209 CTACTGATTTTGGCTCTCAATTCCTCTTCACTCTTCACTGTTCTCTTATTAATGAACGCATCAG 1268
Qy 341 AlaGlnIleAspHisTyrLeuGlyLeuAlaAsnLysAsnValLysAspAlaMetAlaLys 360
Db 1269 GCACAGATAGATCATTTATCTAGGACTTGCATAAAGAAATGTTAAAGATGCTATGGCTAAA 1328
Qy 361 IleGlnAlaLysIleProGlyLeuLysArgLysAlaGlu 373
Db 1329 ATCCAGCAAAATCCCTGGATTGAGCGCAAGCTGAA 1367

RESULT 10
AF148538
LOCUS Homo sapiens 2235 bp mRNA linear PRI 09-SEP-2000
DEFINITION Homo sapiens reticulon 4b mRNA, complete cds.
ACCESSION AF148538
VERSION AF148538.1 GI:10039552
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 2235)
Yang, J., Yu, L., Bi, A.D. and Zhao, S.Y.
Assignment of the human reticulon 4 gene (RTN4) to chromosome
2p14-->2p13 by radiation hybrid mapping
Cytogenet. Cell Genet. 88 (1-2), 101-102 (2000)
JOURNAL
MEDLINE 20237542
PUBMED 10773680
REFERENCE 2 (bases 1 to 2235)
Zhou, Y., Yu, L. and Zhao, S.Y.
Direct Submission
AUTHORS
TITLE Submitted (05-MAY-1999) Lab of Human Gene Research, Institute of
Genetics, Fudan University, No. 220 Handan Rd., Shanghai 200433,
P.R.China
FEATURES
source Location/Qualifiers
1. 2235
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
142..1263
/note="RTN4b"
/codon_start=1
/product="reticulon 4b"
/protein_id="AAG12177.1"
/db_xref="GI:10039553"
/translation="MEDLDQSLVSSSDSPRPQPAFKYQFVREPEDEEEEEEEED
```

```
EDBLELEVLERKPAAGLSAAVPVPTAAGAFIMDFGND FVPAPRGPPLUPAAPVAP
ERQSPWDPSVSTVPAPSLAAAVSPSKLPEDDPPAPPPPPPPASVSPQAEVWT
PPAPAPAPSPSTPAAPKRRGSSGVVDLLYMRDIKKTGVFGASLFLLLSLTVFSIV
SVTAVTALLSLVTSIFRLYKGVIOAKQKSDGHPFRAYLESVALISEELVOKYSNSA
LGHVNCITIKELRRLFLVDDLVDSLKFAVLMWFTYVYVGLFNGLTLLILALISLSPV
IYRHVQAQIDHYLGLANKVVKOMAKIQAIKGLKRAE"
2208..2213
2225
polyA_signal
polyA_site
ORIGIN
Alignment Scores: 5.81e-67 Length: 2235
Pred. No.: 1879.00 Matches: 370
Score: 99.20% Conservative: 0
Percent Similarity: 99.20% Mismatches: 3
Best Local Similarity: 98.38% Indels: 0
Query Match: 98.38% Gaps: 0
DB: 9
US-09-544-776-2 (1-373) x AF148538 (1-2235)
Qy 1 MetGluAspLeuAspGlnSerProLeuValSerSerSerProArgProGln 20
Db 142 ATGGAAGACCTGGACCACTCTCTTGGTCTCTCTCGGACAGCCACCCGCGCGCAG 201
Qy 21 ProAlaPheLysTyrGlnPheValArgGluProGluAspGluGluGluGlu 40
Db 202 CCGCGCTTCAAGTACCAGTTCGTGGAGGAGCCGAGGAGGAGGAGGAGGAG 261
Qy 41 GluGluLysAspGluAspGluLeuGluGluValLeuGluArgLysProAla 60
Db 262 GAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 321
Qy 61 AlaGlyLeuSerAlaAlaProValProThrAlaProAlaAlaGlyAlaProLeuMetAsp 80
Db 322 GCCGCGCTGTCCGCGCGCCAGTGCACCGCCCTCCGCGCGCGCGCGCGCGCG 381
Qy 81 PheGlyAsnAspPheValProAlaProArgGlyPheLeuProAlaAlaProVal 100
Db 382 TTCGGAATGACTTCGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGTC 441
Qy 101 AlaProGluArgGlnProSerTyrAspProSerProValSerThrValProAlaPro 120
Db 442 GCCCGGAGCGGAGCGCGCGCTTGGGAGCCGCGCGCGCGCGCGCGCGCGCGCA 501
Qy 121 SerPheLeuSerAlaAlaValSerProSerLysLeuProGluAspGluProPro 140
Db 502 TCCCGCGCTGTCTGCTCCGCGAGTCTCGCCCTTCAAGCTCCCTGAGGAGCAGCCTCCG 561
Qy 141 AlaArgProProProProProAlaSerValSerProGlnAlaGluProValTyrThr 160
Db 562 GCCCGCGCTCCCGCTCTCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 621
Qy 161 ProProAlaProAlaProAlaProProSerThrProAlaAlaProLysArgArgGly 180
Db 622 CCGCCAGCGCGCGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 681
Qy 181 SerSerGlySerValValValValValValValValValValValValValValVal 200
Db 682 TCCCTCGGCTCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 741
Qy 201 ValPheGlyAlaSerLeuPheLeuLeuSerLeuThrValPheSerIleValSerVal 220
Db 742 GTGTTTGGTGCCAGCCTATTCTCTGCTTTCATTGACAGTATTCAGCATTTGTGAGCGTA 801
Qy 221 ThrAlaTyrIleAlaLeuAlaLeuSerValThrIleSerProArgIleTyrLysGly 240
Db 802 ACAGCCTACATTCGCTTGGCCCTGCTCTCTGTGACCATCAGCTTTAGGATATACAAGGT 861
Qy 241 ValIleGlnAlaIleGlnLysSerAspGluGlyHisProPheArgAlaTyrLeuGluSer 260
Db 862 GTGATCCAGCTATCCAGAAATCAGATGAAGCCACCANTTCAGGCGCATATCTGGAAATCT 921
Qy 261 GluValAlaIleSerGluGluLeuValGlnLysTyrSerAsnSerAlaLeuGlyHisVal 280
```

Db	922	GAAGTTGCTATATCTCAGAGATGGTTCAGAGTACAGTAATCTCTTGGTCAATG	981
Qy	281	AsnCysThrIleLysGluLeuArgArgLeuPheLeuValAspAspLeuValAspSerLeu	300
Db	982	AACTGCACGATAAAGGAACCTCAGCGCCCTCTTCTTAGTTGATGATTTAGTTGATCTCTG	1041
Qy	301	LysPheAlaValLeuMetTrpValPheThrTyrValGlyAlaLeuPheAsnGlyLeuThr	320
Db	1042	AAGTTTGCAGTGTGATGTGGGTATTTACTATGTTGGTGCCTTGTATTAATGGTCTGACA	1101
Qy	321	LeuLeuIleLeuAlaLeuIleSerLeuPheSerValProValIleTyrGluArgHisGln	340
Db	1102	CTACTGATTTTGGCTCTCATTTCACTCTTCAGTGTTCTCTGTTATTAAGACGATCAG	1161
Qy	341	AlaGlnIleAspHisTyrLeuGlyLeuAlaAsnLysValLysAspAlaMetAlaLys	360
Db	1162	GCACAGATAGATCATTTATCTAGGACTTGCAAAATAAGAATGTTAAAGATGCTATGGCTAAA	1221
Qy	361	IleGlnAlaLysIleProGlyLeuLysArgLysAlaGlu	373
Db	1222	ATCCAAGCAAAATCCCTGGATTGAAGCGCAAGCTGAA	1260
RESULT 11			
AF132047			
LOCUS			
DEFINITION Homo sapiens foocen-m mRNA, complete cds.			
ACCESSION AF132047			
VERSION AF132047.1 GI:4838516			
KEYWORDS			
SOURCE Homo sapiens (human)			
ORGANISM Homo sapiens			
REFERENCE 1 (bases 1 to 2276)			
AUTHORS Ito.T. and Schwartz,S.M.			
TITLE Cloning of a member of the reticulon gene family in human:			
JOURNAL ubiquitous type			
REFERENCE 2 (bases 1 to 2276)			
AUTHORS Ito.T. and Schwartz,S.M.			
TITLE Direct Submission			
JOURNAL Submitted (27-FEB-1999) Pathology, University of Washington, 1959			
FEATURES			
source			
1. .2276			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/tissue_type="placenta"			
180. .1301			
/note="member reticulon gene family"			
/codon_start=1			
/product="foocen-m"			
/protein_id="AAD31021.1"			
/db_xref="GI:4838517"			
/translation="MEDLDQSLVSSSDSPRPQPAFKYQFVREPEDEEEEEEED			
EDEDLEELVLRKPAAGLSAAPVTPAAGAPLMDFGNDFVPPAPRGLPAPAPVAP			
ERQPSWDPSSTVPAPSPLSAAAVSPSKLPEDDEPPARPPPPPPASVQPAEPVPT			
PPAPAPAPPTPAAPKRRSGSVVVDLLYWRDIDKKTGWFGASLFLILLSITVSI			
SVPTAYALLSVTISFRYKGYIOAIQKSDGHPFRAYLSEVAISELVQKYSNSA			
LGHVNCTIKELARLFLVDLDVSLKPAVLMWVTVYVGLFNSLILALISLFSVPV			
IYERHQAIQDHYLGLANKNVKDMAKIQIKIPLKRAE"			
ORIGIN			
Alignment Scores:			
Pred. No.:			
Score:			
Percent Similarity:			
Best Local Similarity:			
Query Match:			
DB:			

US-09-544-776-2 (1-373) x AF132047 (1-2276)			
Qy	1	MetGluAspLeuAspGlnSerProLeuValSerSerSerAspSerProProArgProGln	20
Db	180	ATGGAAGACCTGGACAGTCTCTCTGCTCTCGTCTCGACAGCCACCCCGCCGAC	239
Qy	21	ProAlaPheLysTyrGlnPheValArgGluProGluAspGluGluGluGluGlu	40
Db	240	CCCGCGTTCAAGTACAGTTCTGAGGGAGCCGAGGAGGAGGAGGAGGAGGAG	299
Qy	41	GluGluGluAspGluAspGluAspGluGluGluValLeuGluArgLysProAla	60
Db	300	GAAGAGGAGGACGAGGACGAGACCTCGAGGAGCTGGAGTCTGCGAGGAGGAG	359
Qy	61	AlaGlyLeuSerAlaAlaProValProThrAlaProAlaAlaGlyAlaProLeuMetAsp	80
Db	360	GC CGGCGTCTCGCGCGCCAGTGCCTCCGCGCGCGCGCGCGCGCGCGCGCG	419
Qy	81	PheGlyAsnAspPheValProProAlaProAspGlyPheLeuProAlaAlaProProVal	100
Db	420	TTCCGGAATGACTTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	479
Qy	101	AlaProGluArgGlnProSerTrpAspProSerProValSerSerThrValProAlaPro	120
Db	480	GC CGCGGAGCGGCGCGCTTTGGGACCGGACCGCGCGTCTCGACCGTGC CGCG	539
Qy	121	SerPheLeuSerAlaAlaValSerProSerLysLeuProGluAspAspGluProPro	140
Db	540	TC CGCGCTGCTGCTGCGCGCGCTCTCGCGCTCCCAAGCTCCCTGAGGACGAC	599
Qy	141	AlaArgProProProProProProAlaSerValSerProGlnAlaGluProValTrpThr	160
Db	600	GC CGCGCGCTCCCGCTCTCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCG	659
Qy	161	ProProAlaProAlaProAlaProProSerThrProAlaAlaProLysArgArgGly	180
Db	660	CGCGCAGCGCGCGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	719
Qy	181	SerSerGlySerValValValValValValValValValValValValValValVal	200
Db	720	TCCTCGGCGTCACTGGTGTGGTTCCTCTGACCTCTGACCTGAGAGACATTAAGA	779
Qy	201	ValPheGlyAlaSerLeuPheLeuLeuSerLeuThrValPheSerIleValSerVal	220
Db	780	GTGTTTGGTGCAGGCTATTCTCTGCTCTTCATTGACAGTATTTCAGCATTTGAG	839
Qy	221	ThrAlaTyrIleAlaLeuAlaLeuSerValThrIleSerProArgIleTyrLysGly	240
Db	840	ACAGCCTACATTGCTTGGCCCTGCTCTGTGACCATCAGCTTTAGCATATACAA	899
Qy	241	ValIleGlnAlaIleGlnLysSerAspGluGlyHisPropheArgAlaTyrLeuGluSer	260
Db	900	GTGATCCAGCTATCCAGAAATCAGATGAAGGCCACCCATTCAGGCGCATATCT	959
Qy	261	GluValAlaIleSerGluGluLeuValGlnLysTyrSerAsnSerAlaLeuGlyHisVal	280
Db	960	GAAGTTGCTATATCTGAGGAGTTGGTTCAAGAGTACAGTAATCTCTCTTGGT	1019
Qy	281	AsnCysThrIleLysGluLeuArgLeuPheLeuValAspAspLeuValAspSerLeu	300
Db	1020	AACTGCACGATAAAGAACTCAGGGCGCTCTCTTAGTTGATGATTTAGTTAGTT	1079
Qy	301	LysPheAlaValLeuMetTrpValPheThrTyrValGlyAlaLeuPheAsnGlyLeuThr	320
Db	1080	AAGTTTGCAGTGTGATGTGGGTATTTTACCTATGTGGTGCCTTGTATTAATGG	1139
Qy	321	LeuLeuIleLeuAlaLeuIleSerLeuPheSerValProValIleTyrGluArgHisGln	340
Db	1140	CTACTGATTTTGGCTCTCATTTCACTCTCTAGTGTCTCTGTTATTAAGACGCG	1199
Qy	341	AlaGlnIleAspHisTyrLeuGlyLeuAlaAsnLysAsnValLysAspAlaMetAlaLys	360
Db	1200	GCACAGATAGATCATTTATCTAGGACTTGCAATAAGAATGTTAAAGATGCTATG	1259



```
Qy 361 ileGlnAlaYsilleProGlyLeuLysArgLysAlaGlu 373
Db 1260 ATCCAGCAAAATCCCTGGATTGAGCGCAAAAGCTGAA 1298

RESULT 12
AY102277
LOCUS Homo sapiens RTN4 isoform B1 (RTN4) mRNA, complete cds;
DEFINITION alternatively spliced.
ACCESSION AY102277
VERSION AY102277.1 GI:26800561
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2332)
Oertle,T., Huber,C., van der Putten,H. and Schwab,M.E.
Genomic structure and functional characterisation of the promoters
of human and mouse nogo/rtn4
J. Mol. Biol. 325 (2), 299-323 (2003)
23376540
MEDLINE
PUBMED 12488097
REFERENCE
2 (bases 1 to 2332)
Oertle,T. and Schwab,M.E.
Direct Submission
Submitted (07-MAY-2002) Brain Research Institute, University of
Zurich and ETH Zurich, Winterthurerstr.190, Zuerich 8057,
Switzerland
3 (bases 1 to 2332)
Van der Putten,H.
Direct Submission
Submitted (07-MAY-2002) Nervous System Research, Novartis Pharma
Inc., Basel, Switzerland
FEATURES
source
1..2332
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="2"
/map="2p16"
1..2332
/gene="RTN4"
/notes="synonym: NOGO"
1..244
/gene="RTN4"
/evidence=experimental
245..1366
/notes="NOGO-B; RTN4-B1; alternatively spliced"
/codon_start=1
/product="RTN4 isoform B1"
/protein_id="AAM64246.1"
/db_xref="GI:26800562"
/translators="MEDDQSLVSSSDSPRPQPAKQYQVREPDEEEEEEBED
EDDLSELEVERKPAAGLSAAVPPTAPAGAPLMDGNDPVPAPRGPLPAAAPPVAP
ERQSPDPSVSTVPAPSLSAAVSPSKLPEDDPPAPPPPPASVSPQAEVPT
PPAPAPSPSTAAKRGSSGSVVLLYWRDIIKKTGVVFGASLFLLLSLTVFSIV
SVATYALALSVTISFRIYKGVIOAIQKSGDEGHPFRAYLESEVALISEELVQYSNSA
LGHVNTCTIKELRLFLVDLDSLKFAVLMVFTYVFGALPFLGLLILALISLFSVPV
IYERHOAQIDHYLGANKVNDAMAKIQAIPLGRKAE"
1367..2332
/gene="RTN4"

3'UTR
ORIGIN
Alignment Scores:
Pred. No.: 6,05e-67 Length: 2332
Score: 1879.00 Matches: 370
Percent Similarity: 99.20% Conservative: 0
Best Local Similarity: 99.20% Mismatches: 3
Query Match: 98.38% Indels: 0
DB: 9 Gaps: 0
```

```
US-09-544-776-2 (1-373) x AY102277 (1-2332)
Qy 1 MetGluAspLeuAspGlnSerProLeuValSerSerSerAspSerProProArgProGln 20
Db 245 ATGGAAGACCTGGACCAAGTCTCTCTGGTCTCGTCTCGACAGCCACCCGCGCCGAG 304
Qy 21 ProAlaPheLysTyrGlnPheValArgGluProGluAspGluGluGluGluGlu 40
Db 305 CCCGCGTTCAAGTACCAGTTCTGTGAGGAGCCCGAGGAGGAGGAGGAGGAG 364
Qy 41 GluGluGluAspGluAspGluAspLeuGluGluValLeuGluArgLysProAla 60
Db 365 GAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 424
Qy 61 AlaGlyLeuSerAlaAlaProValProThrAlaProAlaAlaGlyAlaProLeuMetAsp 80
Db 425 GCCGGGCTGTCCGGGCCCCAGTCCACCCCTGCCCGCGCGGCCCTCTGATGGAC 484
Qy 81 PheGlyAsnAspPheValProAlaProAlaProArgGlyPheLeuProAlaAlaProVal 100
Db 485 TTCGGAATGACTTCGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 544
Qy 101 AlaProGluArgGlnProSerTrpAspProSerProValSerSerThrValProAlaPro 120
Db 545 GCCCGGAGCGGAGCGCGCTCTTGGGACCCGAGCCCGGTGTCTCGACCGTCCCGGCCA 604
Qy 121 SerPheLeuSerAlaAlaValSerProSerLysLeuProGluAspAspGluProPro 140
Db 605 TCCCCGCTGTCTGCTCGCGAGTCTCGCCCTCCCAAGCTCCTCGAGGACGAGCGCTCCG 664
Qy 141 AlaArgProProProProProAlaSerValSerProGlnAlaGluProValTrpThr 160
Db 665 GCCCGGCTCTCCCTCTCTCCCGCGCGCGGTGAGCCCCCAGGAGCGCGGTGGACC 724
Qy 161 ProProAlaProAlaProAlaProProSerThrProAlaAlaProLysArgArgGly 180
Db 725 CCGCCAGCCCGCGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 784
Qy 181 SerSerGlySerValValValAspLeuLeuLeuLeuLeuLeuLeuLeuLeuVal 200
Db 785 TCCTCGGGCTCAGTGTGTGTGACCTCTCTGTCTGAGAGACATTAAGAAGACATGGAGTG 844
Qy 201 ValPheGlyAlaSerLeuPheLeuLeuLeuSerLeuThrValPheSerIleValSerVal 220
Db 845 GTGTTTGTGGCCAGCCTATTCTCTGCTGCTTTCATTGACAGTATTGACATTGTGAGCGTA 904
Qy 221 ThrAlaTyrIleAlaLeuAlaLeuLeuSerValThrIleSerProArgIleTyrLysGly 240
Db 905 ACAGCCTACATTGCTTGGCCCTGCTCTCTGTGACCATCAGCTTTAGGATATACAAGGGT 964
Qy 241 ValIleGlnAlaIleGlnLysSerAspGluGlyHisProPheArgAlaTyrLeuGluSer 260
Db 965 GTGATCAAGCTATCCAGAAATCAGATGAAGGCCACCCATTACGGGATATCTGGAACTCT 1024
Qy 261 GluValAlaIleSerGluGluLeuValGlnLysTyrSerAsnSerAlaLeuGlyHisVal 280
Db 1025 GAAGTTGTATATCTCAGGAGTTGGTTTCAAGATACAGTAATTTCTGCTCTTGTTCATGTG 1084
Qy 281 AsnCythrIleLysGluLeuArgArgLeuPheLeuValAspAspLeuValAspSerLeu 300
Db 1085 AACTGACGATAAAGGAACCTCAGCGCGCTCTCTTCTAGTTGATGATTTAGTTGTTCTCTG 1144
Qy 301 LysPheAlaValLeuMetTrpValPheThrTyrValGlyValAlaLeuPheAsnGlyLeuThr 320
Db 1145 AAGTTTGCAGTGTGTGATGTGGGTATTTACTATGTTGGTGCTCTGTTTAAATGCTCTGACA 1204
Qy 321 LeuLeuIleLeuAlaLeuIleSerLeuPheSerValProValIleTyrGluArgHisGln 340
Db 1205 CTACTGATTTGGCTCTCTCATTTTCACCTCTTTCAGTGTCTCTGTTTATTTATGACGGCATCAG 1264
Qy 341 AlaGlnIleAspHisTyrLeuGlyLeuAlaAsnLysAsnValLysAspAlaMetAlaLys 360
Db 1264
```

Db 1265 GCACAGATAGATCATTTATCTAGGACTTGCAAAATAAGAAATGTTAAAGATGCTATGGCTAAA 1324

Qy 361 IleGlnAlaLysIleProGlyLeuLysArgLysAlaGlu 373  
 |||||  
 Db 1325 ATCCAAGCAAAATCCCTGGATTGAAGCGCAAGACTGAA 1363  
 |||||

RESULT 13

AK130812 1525 bp mRNA linear PRI 10-SEP-2003

LOCUS Homo sapiens cDNA FLJ27302 fis, clone TMS04776.

DEFINITION AK130812

ACCESSION AK130812

VERSION AK130812.1 GI:34527696

KEYWORDS oligo capping; fis (full insert sequence).

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS Suzuki, O., Sasaki, N., Aotsuka, S., Shoji, T., Ichihara, T., Shihata, N., Matsumoto, K., Hirano, M., Sano, S., Nomura, R., Yoshikawa, Y., Matsumura, Y., Moriya, S., Chiba, E., Momiyama, H., Onogawa, S., Kaeriyama, S., Satoh, N., Matsunawa, H., Takahashi, E., Kataoka, R., Kuga, N., Kuroda, A., Satoh, I., Kamata, K., Takami, S., Terashima, Y., Watanabe, M., Suzuki, Y., Hata, H., Nakagawa, K., Mizuno, S., Morinaga, M., Kawamura, M., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Nishikawa, T., Sugiyama, A., Kawakami, B., Nagai, K., Isogai, T. and Sugano, S.

TITLE NEDO human cDNA sequencing project

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 1525)

AUTHORS Sugano, S. and Suzuki, Y.

TITLE Direct Submission

JOURNAL Submitted (31-JUL-2003) Sumio Sugano, Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure, Human Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan (E-mail: flcdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286, Fax:81-3-5449-5416)

COMMENT NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction and 5'-end one pass sequencing: Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure, Human Genome Center; 3'-end one pass sequencing: RAB; clone selection for full insert sequencing: RAB and Helix Research Institute.

FEATURES

source Location/Qualifiers

1..1525

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="TMS04776"

/cissue\_type="thymus"

/clone\_lib="TMS"

/notes="cloning vector: pMB18SFL3"

ORIGIN

Alignment Scores:

Pred. No.:	5,28e-67	Length:	1525
Score:	1876.00	Matches:	369
Percent Similarity:	99.20%	Conservative:	1
Best Local Similarity:	98.93%	Mismatches:	3
Query Match:	98.22%	Indels:	0
DB:	9	Gaps:	0

US-09-544-776-2 (1-373) x AK130812 (1-1525)

Qy 1 MetGluAspLeuAspGlnSerProLeuValSerSerAspSerProArgProGln 20  
 |||||

Db 60 ATGGAAGACCTGGACCAAGTCTCTCTGGTCTCGTCTCGACAGCCACCCCGCGCGAG 119  
 |||||

Qy 21 ProAlaPheLysTyrGlnPheValArgGluProGluAspGluGluGluGlu 40  
 |||||

Db 120 CCGCGCTTCCAGGTACCAAGTTCGTGAGGAGCCCGAGGACGAGGAAGAGGAGGAG 179  
 |||||

Qy 41 GluGluGluAspGluAspGluAspGluGluGluValLeuGluArgLysProAla 60  
 |||||

Db 180 GAAGAGGAGGACGAGGACGAGACCTGGAGGAGCTGGAGGTCTGGAGGAGGAGCCCGCC 239  
 |||||

Qy 61 AlaGlyLeuSerAlaAlaProValProThrAlaProAlaAlaGlyAlaProLeuMetAsp 80  
 |||||

Db 240 GCGGGGCTGTCGCGGCGCCAGTGCCACCGCCCTGCCGCGCGCGCGCCCTGATGGAC 299  
 |||||

Qy 81 PheGlyAsnAspPheValProProAlaProArgGlyPheLeuProAlaAlaProProVal 100  
 |||||

Db 300 TTTCGAAATGACTTCGTGCGCGCGCGCGCGGACCCCTCGCGCGCTCCCGCCGTC 359  
 |||||

Qy 101 AlaProGluArgGlnProSerTrpAspProSerProValSerSerThrValProAlaPro 120  
 |||||

Db 360 GCGCGGAGCGGACCGCTTTGGGACCGGACCGCGGTGCTGACCGGTGCCCGGCCA 419  
 |||||

Qy 121 SerPheLeuSerAlaAlaAlaValSerProSerLysLeuProGluAspGluProPro 140  
 |||||

Db 420 TCCCGGCTGCTGCTGCTGCCGCGCTCTCGCCCTCCAAGCTCCCTGAGGACGAGCCTCG 479  
 |||||

Qy 141 AlaArgProProProProProProAlaSerValSerProGlnAlaGluProValTrpThr 160  
 |||||

Db 480 GCGCGGCTCTCCCTCTCTCCCGCGCGCGGTGAGCGCCCGCAGGACGAGCCGTGTGGAC 539  
 |||||

Qy 161 ProProAlaProAlaProAlaAlaProProSerThrProAlaAlaProLysArgArgGly 180  
 |||||

Db 540 CCGCAGCGCCCGGCTCCCGCGCGCCCTCCACCCCGCGCGCCCGCAGCGAGGGC 599  
 |||||

Qy 181 SerSerGlySerValValValAspLeuLeuLeuTyrTrpArgAspIleLysLysThrGlyVal 200  
 |||||

Db 600 TCCTCGGCTCAGTGGTTGTTGACCTCTCTGTACTGGAGACATTAAGAAGACTGGAGTG 659  
 |||||

Qy 201 ValPheGlyAlaSerIlePheLeuLeuLeuSerLeuThrValPheSerIleValSerVal 220  
 |||||

Db 660 GTGTTTGGTGGCAGGCTATTCTCGTCTCTTTCATTGACAGATATTGACGATTTGAGCGTA 719  
 |||||

Qy 221 ThrAlaTyrIleAlaLeuAlaLeuLeuSerValThrIleSerProArgIleTyrLysGly 240  
 |||||

Db 720 ACAGCTACATTTGCTTGGCCCTGCTCTCTGTGACCATCAGCTTTAGGATATACAAGGT 779  
 |||||

Qy 241 ValIleGlnAlaIleGlnLysSerAspGluGlyHisProPheArgAlaTyrLeuGluSer 260  
 |||||

Db 780 GTGATCCAAGCTATCCAGAAATCAGATGAAGCCACCCATTGAGGCGATATCTGGAATCT 839  
 |||||

Qy 261 GluValAlaIleSerGluLeuValGlnLysTyrSerAsnSerAlaLeuGlyHisVal 280  
 |||||

Db 840 GAAGTTGCTATATCTGAGGAGTTGTTTCAAGATGACGTAATCTGCTCTTGGTCTGTG 899  
 |||||

Qy 281 AsnCysThrIleLysGluLeuArgArgLeuPheLeuValAspAspLeuValAspSerLeu 300  
 |||||

Db 900 AACTGCACGATNAAGAACTCAGGGCGCTCTTCTTAGTTGATGATTTAGTTGATTTCTCTG 959  
 |||||

Qy 301 LysPheAlaValLeuMetTrpValPheThrTyrValGlyAlaLeuPheAsnGlyLeuThr 320  
 |||||

Db 960 AAGTTTGCAGTGTGATGTGGGTATTTACCTATGTTGGTCCCTTGTATTAATGCTGTACA 1019  
 |||||

Qy 321 LeuLeuIleLeuAlaLeuIleSerLeuPheSerValProValIleTyrGluArgHisGln 340  
 |||||

Db 1020 CTACTGATTTTGGCTCTCATTTTCATCTCTTCACTGTTCCTGTTCTGTTATTATGAACGGCATCAG 1079  
 |||||

Qy 341 AlaGlnIleAspHisTyrLeuGlyLeuAlaAlaAsnLysAsnValLysAspAlaMetAlaLys 360  
 |||||

Db 1080 GCACAGATAGATCATTTACTAGGACTTGCANATAGATGTTAAGATGCTATGGCTAAA 1139  
 |||||

Qy 361 IleGlnAlaLysIleProGlyLeuLysArgLysAlaGlu 373  
 |||||

Db 1140 ATCCAAGCAAAATCCCTGGATTGAAGCGCAAGACTGAA 1178  
 |||||

RESULT 14

LOCUS CQ769577

DEFINITION Sequence 303 from Patent WO2003058021.

ACCESSION CQ769577

LOCUS CQ769577 1599 bp DNA linear PAT 04-MAR-2004



```
VERSION CQ769577.1 GI:45113880
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 Koenig-Hoffman, K., Kazinski, M., Schaefer, R. and Kesper, B.
Novel apoptosis-inducing dna sequences
Patent: WO 2003058021-A 303 17-JUL-2003;
Xantos Biomedicine AG (DE)
FEATURES
source
1. .1599
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 6,06e-67 Length: 1599
Score: 1875.00 Matches: 369
Percent Similarity: 98.93% Conservative: 0
Best Local Similarity: 98.93% Mismatches: 4
Query Match: 98.17% Indels: 0
DB: 6 Gaps: 0
US-09-544-776-2 (1-373) x CQ769577 (1-1599)
Qy 1 MetGluAspLeuAspGlnSerProLeuValSerSerSerAspSerProProArgProGln 20
Db 116 ATGGAAGACCTGGACCAAGTCTCTCTGCTCGTCTCGGACAGCCACCCGCGCGCAG 175
Qy 21 ProAlaPheLysTyrGlnPheValArgGluProGluAspGluGluGluGluGlu 40
Db 176 CCGCGGTTCAGTACCACTTCGTGGAGGACCCCGAGGACGAGGAGGAGGAGGAG 235
Qy 41 GluGluAspGluAspGluAspLeuGluGluGluGluGluGluGluGluGluGlu 60
Db 236 GAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 295
Qy 61 AlaGlyLeuSerAlaAlaProValProThrAlaProAlaAlaGlyAlaProLeuMetAsp 80
Db 296 GCCGGGCTGTCCGGCGGCCCGAGTGCACCCCTGCCCGCGCGCGCGCGCGCGCG 355
Qy 81 PheGlyAsnAspPheValProAlaProAlaProArgGlyPheLeuProAlaAlaProVal 100
Db 356 TTCGGAATCACTTCGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGTC 415
Qy 101 AlaProGluArgGlnProSerTrpAspProSerProValSerSerThrValProAlaPro 120
Db 416 GCCCGGAGCGGAGCGCGTCTTGGGACCGGAGCGCGTGTGTGTGTGTGTGTGTGT 475
Qy 121 SerPheLeuSerAlaAlaValSerProSerLysLeuProGluAspGluProPro 140
Db 476 TCCCGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 535
Qy 141 AlaArgProProProProProProAlaSerValSerProGlnAlaGluProValTrpThr 160
Db 536 GCCCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 595
Qy 161 ProProAlaProAlaProAlaProProSerThrProAlaAlaProLysArgGly 180
Db 596 CCGCAGCCCGCGTCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGGC 655
Qy 181 SerSerGlySerValValAlaAspLeuLeuTyrTrpArgAspLeuLysThrGlyVal 200
Db 656 TCCTCGGCGCTCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 715
Qy 201 ValPheGlyAlaSerLeuPheLeuLeuLeuSerLeuThrValPheSerIleValSerVal 220
Db 716 GTGTTGGTCCGCGCTATTCTCTGCTCTTCAATTGACATATTACGATTGTGAGCGTA 775
Qy 221 ThrAlaTyrIleAlaLeuAlaLeuLeuSerValThrIleSerProArgIleTyrLysGly 240
```

```
|||||
776 ACAGCCTACATGCTTGGCCCTCTCTCTGTGACCATCAGCTTTAGGATATACAAGGT 835
Qy 241 ValIleGlnAlaTleGlnLysSerAspGluGlyHisProPheArgAlaTyrLeuGluSer 260
Db 836 GTGATCCAGCTATCCAGAAATCAGATGAAGGCCACCCATTACGGCATATCTGGAACT 895
Qy 261 GluValAlaIleSerGluGluLeuValGlnLysTyrSerAsnSerAlaLeuGlyHisVal 280
Db 896 GAAGTTGCTATATCTGAGGAGTTGGTTTCAGAACTACAGTAATTTCTCTCTTGTGTCATGTG 955
Qy 281 AsnCythrIleLysGluLeuArgArgLeuPheLeuValAspAspLeuValAspSerLeu 300
Db 956 AACTGACGATAAAGGAACCTCAGCGCCCTCTCTTCTAGTTGATGATTAGTTGATCTCTG 1015
Qy 301 LysPheAlaValLeuMetTrpValPheThrTyrValGlyAlaLeuPheAsnGlyLeuThr 320
Db 1016 AAGTTTGCAGTGTGTGATGGGTATTTACTATGTTGGTGGCTTGTATGCTGTGACA 1075
Qy 321 LeuLeuIleLeuAlaLeuIleSerLeuPheSerValProValIleTyrGluArgHisGln 340
Db 1076 CTACTGATTTTGGCTCTCATTTTCACCTCTTCAGTGTTCCTGTTATTATGAACGGCATCAG 1135
Qy 341 AlaGlnIleAspHisTyrLeuGlyLeuAlaAsnLysAsnValLysAspAlaMetAlaLys 360
Db 1136 GCACAGATAGATCATTTATCTAGGACTTGCAATAAGAATGTTAARGATGCTATGGCTAA 1195
Qy 361 IleGlnAlaLysIleProGlyLeuLysArgLysAlaGlu 373
Db 1196 ATCCAAGCAAAATCCCTGATTTGAAGCGCAAGCTGAA 1234
RESULT 15
LOCUS AB015639 2052 bp mRNA linear PRI 03-SEP-1999
DEFINITION Homo sapiens ASY mRNA, complete cds.
ACCESSION AB015639
VERSION AB015639.1 GI:5821139
KEYWORDS ASY.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 Yutsudo, M.
Isolation of a cell death-inducing gene
Published Only in DataBase (1999)
2 (bases 1 to 2052)
Yutsudo, M.
Direct Submission
Submitted (17-JUN-1998) Masuo Yutsudo, Osaka University, Dept. of
Tumor Virol., Res. Inst. Microb. Dis.; 3-1 Yamadaoka, Suita, Osaka
565-0871, Japan (E-mail:yutsudo@biken.osaka-u.ac.jp.
Tel:81-6-879-8313 Fax:81-6-879-8315)
FEATURES
source
1. .2052
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cell_type="fibroblast"
1. .2052
/gene="ASY"
67. .1188
/gene="ASY"
/codon_start=1
/protein_id="BAA83712.1"
/db_xref="GI:5821140"
/translation="MEDLDQSLVSSDPPRPQPAFKYQFVREPBDEBEDEBEDEBEDE
EDELELEVLERKPAAGLSAAPTAPAGAPLMDFGNDFFVPPRGLPAPAPVAP
ERQPCWDSPSVSTVPAPSLSAVAVPSKLQDDDEPPRPPPPPPPPPPPPPPPPVWT
PAPAPAAPSTPAAPKRRGSGSVVVDLLIYWRDKKTGTVFGASIFLLSLTTFVSI
SVTATLALLSVTISFLYKGVIOAIKQSDGHPRAYLSEVAISELVLKYSNSA
LGHVNCITIKELRLFLVDLVSLLKPAFLMWVFTYVYVGLFNGLLILALISLPSVPV
IYERHQAOIDHLYGLANKNVKQAMAKIQAKIFGLKRAE"
```

## ORIGIN

Alignment Scores: 1.12e-66 Length: 2052  
Pred. No.: 1871.00 Matches: 368  
Score: 98.93% Conservative: 1  
Percent Similarity: 98.93% Mismatches: 4  
Best Local Similarity: 97.96% Indels: 0  
Query Match: 9 Gaps: 0  
DB:

US-09-544-776-2 (1-373) x AB015639 (1-2052)

QY 1 MetGluAspLeuAspGlnSerProLeuValSerSerSerAspSerProProArgProGln 20  
DB 67 ATGGAGAGCTTGGACCAAGTCTCTCTGTGCTCTGTCTCGACAGCCACCCTGGCCGCGAG 126  
QY 21 ProAlaPheLysTyrGlnPheValArgGluProGluAspGluGluGluGlu 40  
DB 127 CCGCGCTTCAAGTACCAGTTCTGTGAGGAGCCGAGGAGGAGGAGGAGGAG 186  
QY 41 GluGluGluAspGluAspGluAspLeuGluLeuGluValLeuGluArgLysProAla 60  
DB 187 GAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 246  
QY 61 AlaGlyLeuSerAlaAlaProValProThrAlaProAlaAlaGlyAlaProLeuMetAsp 80  
DB 247 GCGGGCTGTCCGGGCCCCAGTGCACCGCCCTGCGCGCGCGCCCTGATGGAC 306  
QY 81 PheGlyAsnAspPheValProProAlaProAlaProArgGlyPheLeuProAlaProVal 100  
DB 307 TTCGGAATGACTTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 366  
QY 101 AlaProGluArgGlnProSerTrpAspProSerProValSerSerThrValProAlaPro 120  
DB 367 GCGCGGAGCGGAGCGCGGTGTTGGGACCCGAGCGCGGTCTGTCGACCCCGCGCA 426  
QY 121 SerPheLeuSerAlaAlaValSerProSerLysLeuProGluAspAspGluProPro 140  
DB 427 TCCCCGCTGTCTGCTGCGCGAGTCTCGCCCTCCAGCTCCCTCAGGACGACGAGCTCCG 486  
QY 141 AlaArgProProProProProAlaSerValSerProGlnAlaGluProValTrpThr 160  
DB 487 GCGCGGCTTCCGCTCTCCCGCGGCGAGCGTGGCGCGCGCGCGCGCGCGCG 546  
QY 161 ProProAlaProAlaProAlaProSerThrProAlaAlaProLysArgArgGly 180  
DB 547 CCGCCAGCCCCGCGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 606  
QY 181 SerSerGlySerValValValAspLeuLeuTyrTrpArgAspIleLysLysThrGlyVal 200  
DB 607 TCCTCGGGCTCAGTGGTGTCTGACCTCTCTGTACTGGAGAGACATTAAAGAGACTGGAGTG 666  
QY 201 ValPheGlyAlaSerLeuPheLeuLeuLeuSerLeuThrValPheSerIleValSerVal 220  
DB 667 GTGTTTGGTGGCAGCGCTATTCTCTGTCTGCTTTCAATTGACAGTATTGAGCATTTGAGCGGTA 726  
QY 221 ThrAlaTyrIleAlaLeuAlaLeuLeuSerValThrIleSerProArgIleTyrLysGly 240  
DB 727 ACAGCTCATATTGCTTGGCCCTGCTCTCTGTGACCATCAGCTTTAGGATATACAGGGGT 786  
QY 241 ValIleGlnAlaIleGlnLysSerAspGluGlyHisProPheArgAlaTyrLeuGluSer 260  
DB 787 GTGATCCAGCTATCCAGAAATCAGATGAAGGCCACCCATTTCAGGGCATATCTGGAACTCT 846  
QY 261 GluValAlaIleSerGluIleValGlnLysTyrSerAsnSerAlaLeuGlyHisVal 280  
DB 847 GAAGTTGCTTATATCTGAGGAGTTGGTTTCAGAAAGTACAGTAATCTCTCTCTGCTCATGTG 906  
QY 281 AsnCysThrIleLysGlnLeuArgArgLeuPheLeuValAspAspLeuValAspSerLeu 300  
DB 907 AACTGCACGATTAAGGAATCTAGGCGCCCTCTCTTTAGTTGATGATTTAGTTGATTTCTCTG 966  
QY 301 LysPheAlaValLeuMetTrpValPheThrTyrValGlyAlaLeuPheAsnGlyLeuThr 320

DB 967 AAGTTTGCAGTCTTGTATGTGGGTATTACCTATGTGTGGCTTGTTTAATGGTCTGACA 1026  
QY 321 LeuLeuIleLeuAlaLeuIleSerLeuPheSerValProValIleTyrGluArgHisGln 340  
DB 1027 CTACTGATTTTGGCTCTCATTTTCACTTTCAGTGTCTCTGTATTATTATGAACGGCATCAG 1086  
QY 341 AlaGlnIleAspHisTyrIleuGlyLeuAlaAsnLysAsnValLysAspAlaMetAlaLys 360  
DB 1087 GCACAGATAGATCATTTATCTAGGACTTGCAATATAGAAATGTTAAAGATGCTATGGCTAAA 1146  
QY 361 IleGlnAlaLysIleProGlyLeuLysArgLysAlaGlu 373  
DB 1147 ATCCAAGCAAAATTCCTGGATTGAAGCCCAAGGCTGAA 1185

Search completed: August 3, 2005, 23:12:11  
Job time : 4346 secs

Result No.	Query	Score	Match	Length	DB	ID	Description	
C	1	37.4	1.7	48	3	US-09-580-923-34	Sequence 34, Appl	
	2	37.4	1.7	50	3	US-08-860-038-17	Sequence 17, Appl	
	3	37.4	1.7	50	3	US-09-580-923-17	Sequence 17, Appl	
	4	35.2	1.6	48	3	US-08-979-608A-36	Sequence 36, Appl	
	5	35.2	1.6	48	4	US-09-517-849-36	Sequence 36, Appl	
	6	35.2	1.6	48	4	US-09-616-289-36	Sequence 36, Appl	
	7	31	1.4	39	1	US-08-068-747-9	Sequence 9, Appl	
	8	28	1.2	45	3	US-08-979-608A-30	Sequence 30, Appl	
	9	28	1.2	45	4	US-09-517-849-30	Sequence 30, Appl	
	10	28	1.2	45	4	US-09-616-289-30	Sequence 30, Appl	
C	11	27.6	1.2	45	2	US-08-356-786-13	Sequence 13, Appl	
	12	25.8	1.2	50	4	US-09-513-999C-30835	Sequence 30835, A	
	13	25.6	1.1	50	4	US-08-956-171E-2153	Sequence 2153, Ap	
	14	25.6	1.1	50	4	US-08-781-986A-2153	Sequence 2153, Ap	
	15	25.2	1.1	30	1	US-08-068-747-4	Sequence 4, Appl	
	16	25.2	1.1	41	3	US-08-589-109A-12	Sequence 12, Appl	
	17	25.2	1.1	41	3	US-08-113-646A-39	Sequence 39, Appl	
	18	25.2	1.1	42	2	US-08-776-94A-4	Sequence 4, Appl	
	19	25.2	1.1	44	1	US-08-113-646A-40	Sequence 40, Appl	
	20	24.8	1.1	36	1	US-08-113-646A-6	Sequence 6, Appl	
C	21	24.6	1.1	50	3	US-09-371-489-4	Sequence 4, Appl	
	22	24.4	1.1	42	1	US-08-068-747-10	Sequence 10, Appl	
	23	24.4	1.1	50	1	US-08-420-443-1	Sequence 1, Appl	
	24	24.4	1.1	50	4	US-09-621-976-14715	Sequence 14715, A	
	25	24.2	1.1	37	1	US-08-113-646A-38	Sequence 38, Appl	
	26	24.2	1.1	40	3	US-09-306-290-12	Sequence 12, Appl	
	27	24.2	1.1	50	3	US-09-930-181-14	Sequence 14, Appl	



```
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/616,289
; CURRENT FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US/09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US/08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US/60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US/60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PastSeq for Windows Version 4.0
; SEQ ID NO 36
; LENGTH: 48
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-616-289-35

Query Match      1.6%; Score 35.2; DB 4; Length 48;
Best Local Similarity 83.3%; Pred. No. 41;
Matches 40; Conservative 0; Mismatches 8; Indels 0

QY   234 GAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGCAGAC 281
Db    1 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGCAGAC 348

RESULT 7
US-08-068-747-9/c
; Sequence 9, Application US/08068747
; Patent No. 5695933
; GENERAL INFORMATION:
; APPLICANT: Schalling, Martin
; APPLICANT: Hudson, Thomas J.
; APPLICANT: Housman, David E.
; TITLE OF INVENTION: Direct Determination of Expanded
; TITLE OF INVENTION: Nucleotide Repeats in the Human Genome
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/068,747
; FILING DATE: 28-MAY-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: MIT-6141
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 39 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "Synthetic"
US-08-068-747-9

Query Match      1.4%; Score 31; DB 1; Length 39;
```





```
/
/ FILING DATE: January 5, 1996
/ APPLICATION NUMBER: 08/781,986
/ FILING DATE: January 3, 1997
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Mark J. Hyman
/ REGISTRATION NUMBER: 46,789
/ REFERENCE/DOCKET NUMBER: PB248P1
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (240) 314-1224
/ TELEFAX: (301) 309-8439
/ INFORMATION FOR SEQ ID NO: 2153:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 50 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: linear
/ SEQUENCE DESCRIPTION: SEQ ID NO: 2153:
US-08-956-171E-2153
Query Match 1.1%; Score 25.6; DB 4; Length 50;
Best Local Similarity 69.4%; Pred. No. 8.8e+03;
Matches 34; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 617 CCGCGCAGCCCGGCTCCCGCGCGCCCTCCACCGCGCGCGCC 665
Db 1 CCCCCCTCCCGCTCCCGCNCNCCCCCCCCCCCCCCCCCCCC 49

RESULT 14
US-08-781-986A-2153
/ Sequence 2153, Application US/08781986A
/ Patent No. 6737248
/ GENERAL INFORMATION:
/ APPLICANT: Charles Kunsch
/ TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
/ NUMBER OF SEQUENCES: 5255
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Human Genome Sciences, Inc.
/ STREET: 9410 Key West Avenue
/ CITY: Rockville
/ STATE: Maryland
/ COUNTRY: USA
/ ZIP: 20850
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
/ COMPUTER: HP Vectra 486/33
/ OPERATING SYSTEM: MSDOS version 6.2
/ SOFTWARE: ASCII Text
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/781,986A
/ FILING DATE:
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER:
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Benson, Bob
/ REGISTRATION NUMBER: 30,446
/ REFERENCE/DOCKET NUMBER: PB248PP
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (301) 309-8504
/ TELEFAX: (301) 309-8512
/ INFORMATION FOR SEQ ID NO: 2153:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 50 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: linear
US-08-781-986A-2153
Query Match 1.1%; Score 25.6; DB 4; Length 50;
Best Local Similarity 69.4%; Pred. No. 8.8e+03;
Matches 34; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
```

```
QY 617 CCGCGCAGCCCGGCTCCCGCGCGCCCTCCACCGCGCGCGCC 665
Db 1 CCCCCCTCCCGCTCCCGCNCNCCCCCCCCCCCCCCCCCCCC 49

RESULT 15
US-08-068-747-4
/ Sequence 4, Application US/08068747
/ Patent No. 5695933
/ GENERAL INFORMATION:
/ APPLICANT: Schalling, Martin
/ APPLICANT: Hudson, Thomas J.
/ APPLICANT: Housman, David E.
/ TITLE OF INVENTION: Direct Determination of Expanded
/ TITLE OF INVENTION: Nucleotide Repeats in the Human Genome
/ NUMBER OF SEQUENCES: 11
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
/ STREET: Two Militia Drive
/ CITY: Lexington
/ STATE: Massachusetts
/ COUNTRY: USA
/ ZIP: 02173
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/068,747
/ FILING DATE: 28-MAY-1993
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Granahan, Patricia
/ REGISTRATION NUMBER: 32,227
/ REFERENCE/DOCKET NUMBER: MIT-6141
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 617-861-6240
/ TELEFAX: 617-861-9540
/ INFORMATION FOR SEQ ID NO: 4:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 30 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: other nucleic acid
/ DESCRIPTION: /desc = "Synthetic"
US-08-068-747-4
Query Match 1.1%; Score 25.2; DB 1; Length 30;
Best Local Similarity 90.0%; Pred. No. 8.1e+03;
Matches 27; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 238 AGGAGGAGAGAGAGGAGGAGGAGGAGGAGG 267
Db 1 AGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 30

Search completed: August 4, 2005, 02:56:42
Job time : 378 secs
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 3, 2005, 17:13:52 ; Search time 573 Seconds  
(without alignments)  
3853.514 Million cell updates/sec

Title: US-09-544-776-2

Perfect score: 1910

Sequence: 1 MEDLDQSLVSSSDSPRPQ.....VKDAMAKIOAKIPGLKRKAE 373

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp  
-Q=/cgn2\_1/USPTO\_spool\_p/US09544776/runat\_03082005\_181344\_28194/app\_query.fasta\_1.519  
-DB=N Geneseq 16Dec04 -QFWT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPEL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFWT=ptio -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09544776 @CGN 1.1 708 @runat\_03082005\_181344\_28194 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- 1: Geneseq 16Dec04:\*
- 2: Geneseq 1980s:\*
- 3: Geneseq 1990s:\*
- 4: Geneseq 2000s:\*
- 5: Geneseq 2001bs:\*
- 6: Geneseq 2002bs:\*
- 7: Geneseq 2003bs:\*
- 8: Geneseq 2003bs:\*
- 9: Geneseq 2003bs:\*
- 10: Geneseq 2003cs:\*
- 11: Geneseq 2003ds:\*
- 12: Geneseq 2004as:\*
- 13: Geneseq 2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1879	98.4	1122	3	Aaz56888 Human MAG
2	1879	98.4	1122	4	Aaf90325 Human NOG
3	1879	98.4	1216	6	Aba05903 Human RTN
4	1879	98.4	1610	3	Aaz36230 cDNA enco
5	1879	98.4	2235	6	Abv94681 Human pan

RESULT 1																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																					
----------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

Aac64406 Human Nog  
Adi62860 Human apo  
Abk90133 DNA enco  
Aak94408 Human ful  
Adi11137 Full leng  
Aad08386 Human sec  
Aad58284 Rat foocce  
Aaz56886 Human MAG  
Aaf90324 Human NOG  
Abk90134 DNA enco  
Abn86601 Human neu  
Aad07886 Human pol  
Adi13965 Human NOG  
Aad94550 Human NOG  
Aas09453 Human cDN  
Acc81048 Human Nog  
Abv94680 Human pan  
Adg32772 Human rta  
Adi83534 Human dta  
Aaa23454 cDNA enco  
Aab70449 Human bon  
Abx34563 Human mdd  
Adi6433 Construct  
Adp45571 Rat NogoA  
Aad01173 Rat neuro  
Abn86600 Rat neuro  
Aad07888 Mouse pol  
Aaf32725 Human sec  
Adi6433 Nucleotid  
Aad01174 Bovine neu  
Aai98079 Human neu  
Aaf90323 Human NOG  
Aaa72983 Human NSP  
Aav23695 Human NSP  
Aax04379 Human sec  
Adk14166 Human aut  
Aax97587 Extended  
Adp18854 Human sec  
Aav30920 Human sec



XX New polypeptide designated NOGO-C is a splice variant of the human NOGO  
PT gene and may be useful in the treatment of neural disorders including  
PT Alzheimer's and Parkinson's diseases.  
XX  
PS Disclosure; Page 27; 25pp; English.  
XX  
CC The present sequence is that of cDNA encoding human NOGO-B (see  
CC AAB82350). NOGO-B is a previously known splice variant of the human NOGO  
CC gene on chromosome 2p21. The invention relates to a novel NOGO gene  
CC splice variant, NOGO-C (see AAF90323). It provides NOGO-C polypeptides  
CC and polynucleotides, and methods for producing such polypeptides by  
CC recombinant techniques. Also disclosed are methods for utilizing NOGO-C  
CC polypeptides and polynucleotides in the treatment of diseases including  
CC neuropathies, spinal injury, brain injury, stroke, neuronal degeneration,  
CC for example Alzheimer's disease and Parkinson's disease, neuromuscular  
CC disorders, psychiatric disorders and developmental disorders. Also  
CC provided are methods for identifying agonists and antagonists for use in  
CC treating conditions associated with NOGO-C imbalance, and diagnostic  
CC assays for detecting diseases associated with inappropriate NOGO-C  
CC activity or levels  
XX  
SQ Sequence 1122 BP; 224 A; 339 C; 316 G; 243 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 3,01e-75 Length: 1122  
Score: 1879.00 Matches: 370  
Percent Similarity: 99.20% Conservative: 0  
Best Local Similarity: 99.20% Mismatches: 3  
Query Match: 98.38% Indels: 0  
DB: 4 Gaps: 0

US-09-544-776-2 (1-373) x AAF90325 (1-1122)

QY 1 MetGluAspLeuAspGlnSerProLeuValSerSerSerAspSerProProArgProGln 20  
DB 1 ATCGAGACCTGGACACCATCTCTCTGCTCGGTCTCGGACAGCCACCCCGCGCGCAG 60

QY 21 ProAlaPheLysTyrGlnPheValArgGluProGluAspGluGluGluGluGlu 40  
DB 61 CCCGCGTTCAAGTACCAAGTTCTGTGAGGAGCCCGGAGCAGGAGGAGGAGGAGGAG 120

QY 41 GluGluGluAspGluAspGluAspGluGluGluGluGluGluGluGluGluGlu 60  
DB 121 GAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 180

QY 61 AlaGlyLeuSerAlaAlaProValProThrAlaProAlaAlaGlyAlaProLeuMetAsp 80  
DB 181 GCCGGGCTGTCCGGCGCCCGGAGTCCCGCCCGGAGTCCCGCCCGGAGTCCCGGAG 240

QY 81 PheGlyValAsnAspPheValProProAlaProAlaProAlaProAlaProVal 100  
DB 241 TTGCGAAATGACTTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300

QY 101 AlaProGluArgGlnProSerTrpAspProSerProValSerSerThrValProAlaPro 120  
DB 301 GCCCGGAGCGGAGCGGCTCTTGGAGCCCGGAGCCCGGAGTCCCGGAGCCCGGAG 360

QY 121 SerPheLeuSerAlaAlaAlaValSerProSerLysLeuProGluAspGluProPro 140  
DB 361 TCCCGCGTGTCTGTCGCGCGAGTCTCGCCCTCCAAAGCTCCCTGAGGACGAGCCTCCG 420

QY 141 AlaArgProProProProProProAlaSerValSerProGlnAlaGluProValTrpThr 160  
DB 421 GCCCGGCTCTCCCTCTCCCGCGGCGAGGCTGAGCCCCCAGGAGGAGCCCGTGTGAG 480

QY 161 ProProAlaProAlaProAlaProProSerThrProAlaProLysArgArgGly 180  
DB 481 CCGCCAGCCCGGCTCCCGCGCGCGCCCTCCACCCCGCGCGCCCGGAGCGAGCGG 540

QY 181 SerSerGlySerValValValAspLeuLeuTyrTrpArgAspIleLysThrGlyVal 200  
DB 541 TCCTCGGGCTCAGTGGTGTGTGACCTCTGCTACTGGAGAGACATTAAAGAGCTGGAGTG 600

QY 201 ValPheGlyValAspSerLeuPheLeuLeuLeuSerLeuThrValPheSerIleValSerVal 220  
DB 601 GTGTTTGGTGGCCAGCCTATTCTCTGCTGCTTTCATTGACAGTATTTCAGCATTTGTGAGCGTA 660

QY 221 ThrAlaTyrIleAlaLeuAlaLeuLeuSerValThrIleSerProArgIleTyrLysGly 240  
DB 661 ACAGCCCTACATTGCTTGGCCCTCTCTGTGACCATCAGCTTTAGGATATACAAGGGT 720

QY 241 ValIleGlnAlaIleGlnLysSerAspGluGlyHisProPheArgAlaTyrLeuGluSer 260  
DB 721 GTGATCCAGCATTCAGAAATCAGATGAAGGCCACCCATTAGGGCATATCTGGAATCT 780

QY 261 GluValAlaIleSerGluGluLeuValGlnLysTyrSerAsnSerAlaLeuGlyHisVal 280  
DB 781 GAAGTTGTTATATCTCAGGAGTTGGTTTCAGAGTACAGTAATTTCTGCTCTTGTGTCATGTG 840

QY 281 AsnCysThrIleLysGluLeuArgLeuPheLeuValAspAspLeuValAspSerLeu 300  
DB 841 AACTGCAGCATAAAGGAACCTCAGGCGCCTCTCTTAGTTGATGATTTAGTTGATTTCTCTG 900

QY 301 LysPheAlaValLeuMetTrpValPheThrTyrValGlyAlaLeuPheAsnGlyLeuThr 320  
DB 901 AGTTTGCAGTGTGTGATGTGGGTATTACCTATGTTGGTGCCTTGTTTAAATGTTCTGACA 960

QY 321 LeuLeuIleLeuAlaLeuIleSerLeuPheSerValProValIleTyrGluArgHisGln 340  
DB 961 CTACTGATTTGGCT 1020

QY 341 AlaGlnIleAspHisTyrLeuGlyLeuAlaAsnLysAsnValLysAspAlaMetAlaLys 360  
DB 1021 GCACAGATAGATCATTTATCTAGGACTTGCAAAATGAAGATGTTAAAGATGCTATGGCTAAA 1080

QY 361 IleGlnAlaLysIleProGlyLeuLysArgLysAlaGlu 373  
DB 1081 ATCCAAGCAAAATCCCTGATTTGAAGCGCAAGCTGAA 1119

RESULT 3  
ABA05903  
ID ABA05903 standard; cDNA; 1216 BP.  
XX ABA05903;  
AC ABA05903;  
XX  
DT 04-MAR-2002 (first entry)  
XX  
DE Human RTN4B encoding cDNA SEQ ID NO 3.  
XX  
KW Human; RTN4B; gene; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 5..1126  
FT /\*tag= a  
FT /product= "RTN4B"  
XX  
XX CN1311439-A.  
XX  
PN 05-SEP-2001.  
PD  
XX 02-MAR-2000; 2000CN-00111791.  
PF  
XX 02-MAR-2000; 2000CN-00111791.  
PR  
XX (UYFU-) UNIV FUDAN.  
PA  
XX Yu L, Fu Q, Zhao Y;  
PI WPI; 2002-049934/07.  
DR P-PSDB; AAM47954.  
XX  
XX Human RTN 4B protein and coding sequence, its preparation and use.  
XX

PS Claim 9; Page 20 (Disclosure); 27pp; Chinese.

CC The invention relates to human RTN4B protein and coding sequence, useful for providing a cDNA sequence of human RTN4B. The protein is an isomer of CC RTN4 with RTN family members. The present invention also refers to CC polypeptide encoded by a nucleotide sequence and manufacturing method and CC application of the polypeptide and polynucleotide

XX SQ Sequence 1216 BP; 256 A; 354 C; 341 G; 265 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	3-23e-75	Length:	1216
Score:	1879.00	Matches:	370
Percent Similarity:	99.20%	Conservative:	0
Best Local Similarity:	99.20%	Mismatches:	3
Query Match:	98.38%	Indels:	0
DB:	6	Gaps:	0

US-09-544-776-2 (1-373) x ABA05903 (1-1216)

QY 1 MetGluAspLeuAspGlnSerProLeuValSerSerSerSerProProArgProGln 20

DB 5 ATGGAAGACCTGGACGAGTCTCTCTGCTCTGCTCGACAGCCACCCCGCGCGAG 64

QY 21 ProAlaPheLysTyrGlnPheValArgGluProGluAspGluGluGluGluGlu 40

DB 65 CCGCGCTTCAATGACGATTCGTGAGGAGCCCGACGAGCAGGAGGAGGAGGAG 124

QY 41 GluGluGluAspGluAspGluAspGluGluGluGluGluGluGluGluGluGlu 60

DB 125 GAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 184

QY 61 AlaGlyLeuSerAlaAlaProValProThrAlaProAlaAlaGlyAlaProLeuMetAsp 80

DB 185 GCCGGGCTGTCCGCGCCCGCAGTCCCGCCCGCCCGCCCGCCCGCCCGCCG 244

QY 81 PheGlyAsnAspPheValProAlaProAlaProAlaProAlaProAlaProAla 100

DB 245 TTCGGAATGACTTCGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 304

QY 101 AlaProGluArgGlnProSerTrpAspProSerProValSerSerThrValProAlaPro 120

DB 305 GCCCGGAGCGGAGCGCGCTTGGGACCGGCGCGCGCGCGCGCGCGCGCGCG 364

QY 121 SerPheLeuSerAlaAlaValSerProSerLysLeuProGluAspGluProPro 140

DB 365 TCCCCGCTGTCTGCTCCGCGAGTCTCGCCCTCCAGCTCCCTGAGACGACGAGCTCG 424

QY 141 AlaArgProProProProProAlaSerValSerProGlnAlaGluProValTrpThr 160

DB 425 GCCCGGCTTCCCTCTCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 484

QY 161 ProProAlaProAlaProAlaProProSerThrProAlaAlaProLysArgArgGly 180

DB 485 CCGCCAGCCCGGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 544

QY 181 SerSerGlySerValValValAspLeuLeuTyrTrpArgAspIleLysLysThrGlyVal 200

DB 545 TCTCTGGGCTCAGTGTGTGTGACCTCTCTGTTGAGAGACATTAAGAGAGCTGGAGTG 604

QY 201 ValPheGlyAlaSerLeuPheLeuLeuLeuSerLeuThrValPheSerIleValSerVal 220

DB 605 GTGTTTGGTGGCAGCTATTCTCTGCTGCTTTCATTGACAGTATTACAGCTTGTGACGTA 664

QY 221 ThrAlaTyrIleAlaLeuAlaLeuLeuSerValThrIleSerProArgIleTyrLysGly 240

DB 665 ACAGCCTACATTGCCCTTGGCCCTGCTCTGTGACCATCAGCTTTAGGATATACAGGCT 724

QY 241 ValIleGlnAlaIleGlnLysSerAspGluGlyHisProPheArgAlaTyrLeuGluSer 260

DB 725 GTGATCAAGCTATCCAGAAATCAGATGAAGGCCACCCATTACAGGCAATCTCGGAATCT 784

QY 261 GluValAlaIleSerGluGluLeuValGlnLysTyrSerAsnSerAlaLeuGlyHisVal 280

DB 785 GAAGTTGCTATATCTGAGGAGTTGGTTCAGAAGTACAGTAATCTGCTCTTGGTCATGTG 844

QY 281 AsnCysThrIleGlyGluLeuArgLeuPheLeuValAspLeuValAspSerLeu 300

DB 845 AACTGCACGATAAAGAACTCAGGGCGCTCTCTTCTAGTTGATGATTTAGTATTCTCTG 904

QY 301 LysPheAlaValLeuMetTrpValPheThrTyrValGlyAlaLeuPheAsnGlyLeuThr 320

DB 905 AAGTTTGCAGTGTGATGTGGGTATTTACCTATGTGGTCTGTTTAAATGGTCTGACA 964

QY 321 LeuLeuIleLeuAlaLeuIleSerLeuPheSerValProValIleTyrGluArgHisGln 340

DB 965 CTACTGATTTTGGCTCTCATTTCACTCTTCAGTGTCTGTTATTTATGAAAGGCAATCAG 1024

QY 341 AlaGlnIleAspHisTyrIleuGlyLeuAlaAsnLysAsnValLysAspAlaMetAlaLys 360

DB 1025 GCACAGATAGATCAATTCATCTAGGACTTGCAATAAGATGTTAAAGATGCTATGGCTAAA 1084

QY 361 IleGlnAlaLysIleProGlyLeuLysArgLysAlaGlu 373

DB 1085 ATCCAAGCAAAATCCCTGGATTGAAGCGCAAGCTGAA 1123

RESULT 4

AAZ36230

ID AAZ36230 standard; cDNA; 1610 BP.

XX AAZ36230;

XX 22-FEB-2000 (first entry)

DE cDNA encoding a bone marrow secreted protein designated BMS112.

XX Bone marrow secreted protein; bone marrow stromal cell; cytokine; cell proliferation; cell differentiation; hematopoiesis; anaemia; myeloid cell deficiency; lymphoid cell deficiency; myeloid cell; erythroid progenitor cell; colony stimulating factor; granulocyte; monocyte; macrophage; myelo-suppression; megakaryocyte; platelet; platelet disorder; thrombocytopenia; hematopoietic stem cell; stem cell disorder; aplastic anaemia; bone differentiation; paroxysmal nocturnal hemoglobinuria; bone growth; cartilage; tendon; ligament; nerve; wound healing; tissue repair; burn; incision; ulcer; bone fracture; cartilage damage; artificial joint; ss.

OS Homo sapiens.

XX Key Location/Qualifiers

FT CDS 132..1253

FT /\*tag= a

FT /product= "bone marrow secreted protein"

FT polyA\_signal 1516..1521

FT /\*tag= b

XX W09933979-A2.

PN 08-JUL-1999.

PD 18-DEC-1998; 98WO-US027008.

XX 30-DEC-1997; 97US-0068958P.

PR 24-SEP-1998; 98US-0101603P.

PR 30-SEP-1998; 98US-0102540P.

XX (CHIR ) CHIRON CORP.

PI Lin H, Cao L;

XX WPI: 2000-038344/03.

DR P-PSDB; AAY53624.

XX New isolated human polynucleotide and secreted proteins can induce production of other cytokines in certain cell populations.

Db	612	CCGCCAGCCCGGGCTCCCGCGCGCCCTCCACCCCGCGCGCCCAAGCGAGCGGGC	671
Qy	181	SerSerGlySerValValValAspLeuLeuTyrTrpArgAspIleYsLysThrGlyVal	200
Db	672	TCCTCGGGCTCAGTGGTGTGGACCTCTCTGCTGAGAGACATTAAGAAGACTGGAGTG	731
Qy	201	ValPheGlyAlaSerLeuPheLeuLeuLeuSerLeuThrValPheSerIleValSerVal	220
Db	732	GTGTGTTGGTCCAGCCCTATTCTCTGCTGCTTTCATTGACAGTATTCAGCATTTGTGAGCGTA	791
Qy	221	ThrAlaTyrIleAlaLeuAlaLeuLeuSerValThrIleSerProArgIleTyrLysGly	240
Db	792	ACAGCCTACATTGGCTTGGCCCTGCTCTCTGTGACCATGAGCTTTAGGATATACNAGGGT	851
Qy	241	ValIleGlnAlaIleGlnLysSerAspGluGlyHisProPheArgAlaTyrLeuGluSer	260
Db	852	GTGATCCAGCTATCCAGANAATCAGATGAAGGCCACCCATTACGGGCATATCTGGAATCT	911
Qy	261	GluValAlaIleSerGluGluLeuValGlnLysTyrSerAsnSerAlaLeuGlyHisVal	280
Db	912	GAAGTTGCTATATCTCAGGAGTTGGTTTCAGAAAGTACAGTAATTTCTGCTCTTGGTCAATGTC	971
Qy	281	AsnCysThrIleLysGluLeuArgArgLeuPheLeuValAspPheLeuValAspSerLeu	300
Db	972	AATGTCACGATAAAGGAACCTCAGCGGCTCTCTTAGTTGATGATTAGTTGATCTCTCTG	1031
Qy	301	LysPheAlaValLeuMetTrpValPheThrTyrValGlyAlaLeuPheAsnGlyLeuThr	320
Db	1032	AGTTTGCAGTGTGATGGGTATTACCTATGTTGGTGCTTGTTTAATGGTCTGACA	1091
Qy	321	LeuLeuIleLeuAlaLeuIleSerLeuPheSerValProValIleTyrGluArgHisGln	340
Db	1092	CTACTGATTTTGGCTCTCTCAATTCATCTTCAGTGTTCCTGTATTTATGAACGGCATCAG	1151
Qy	341	AlaGlnIleAspHisTyrLeuGlyLeuAlaAsnLysAsnValLysAspAlaMetAlaLys	360
Db	1152	GCAAGATAGATCATTTATCTAGGACTTGGCAATTAAGAATCTTAAAGATGCTATGGCTAAA	1211
Qy	361	IleGlnAlaLysIleProGlyLeuLysArgLysAlaGlu	373
Db	1212	ATCCAGCAAAAATCCCTGGATTGAAGCGCAAGCTGAA	1250
RESULT 5			
ABV94681	ID	ABV94681 standard; cDNA; 2235 BP.	
XX	AC	ABV94681;	
XX	DT		
XX	XX	14-JAN-2003 (first entry)	
XX	XX	Human pancreatic cancer expressed cDNA SEQ ID NO 54.	
KW	XX	Human; pancreas; cancer; gene therapy; vaccine; immunostimulant;	
KW	XX	cytostatic; tumour; gene; ss.	
OS	XX	Homo sapiens.	
XX	XX		
XX	XX	WO200260317-A2.	
XX	XX		
XX	XX	08-AUG-2002.	
XX	XX		
XX	XX	30-JAN-2002; 2002WO-US002781.	
XX	XX		
XX	XX	30-JAN-2001; 2001US-0265305P.	
XX	XX	31-JAN-2001; 2001US-0265682P.	
XX	XX	09-FEB-2001; 2001US-0267568P.	
XX	XX	21-MAR-2001; 2001US-0278651P.	
XX	XX	28-APR-2001; 2001US-0287112P.	
XX	XX	16-MAY-2001; 2001US-0291631P.	
XX	XX	12-JUL-2001; 2001US-0305484P.	
XX	XX	20-AUG-2001; 2001US-0313999P.	
XX	XX	27-NOV-2001; 2001US-0333626P.	

RESULT 5	
ABV94681	
ID	ABV94681 standard; cDNA; 2235 BP.
XX	
XX	ABV94681;
XX	
XX	
DT	14-JAN-2003 (first entry)
XX	
XX	Human pancreatic cancer expressed cDNA SEQ ID NO 54.
XX	
KW	Human; pancreas; cancer; gene therapy; vaccine; immunostimulant;
KW	cytostatic; tumour; gene; ss.
XX	
OS	Homo sapiens.
XX	
XX	WO200260317-A2.
XX	
PD	08-AUG-2002.
XX	
XX	
PF	30-JAN-2002; 2002WO-US002781.
XX	
XX	30-JAN-2001; 2001US-0265305P.
PR	31-JAN-2001; 2001US-0265682P.
PR	09-FEB-2001; 2001US-0267568P.
PR	21-MAR-2001; 2001US-0278651P.
PR	28-APR-2001; 2001US-0287112P.
PR	16-MAY-2001; 2001US-0291631P.
PR	12-JUL-2001; 2001US-0305484P.
PR	20-AUG-2001; 2001US-0313999P.
PR	27-NOV-2001; 2001US-0333626P.
XX	

PA (CORI-) CORIXA CORP.  
 XX Benson DR, Kalos MD, Lodes MJ, Persing DH, Hepler WT, Jiang Y;  
 XX WPI; 2002-627435/67.  
 DR P-PSDB; ABP68601.  
 XX  
 PT New isolated polynucleotide and pancreatic tumor polypeptides, useful for  
 PT diagnosing, preventing and/or treating cancer, particularly pancreatic  
 PT cancer.  
 XX  
 PS Claim 1; SEQ ID NO 54; 300pp + Sequence Listing; English.  
 XX  
 CC The invention relates to an isolated polynucleotide (I) comprising: (a)  
 CC any of a group of over 4000 nucleotide sequences (ABV94628-ABV94145); (b)  
 CC complements of (a); (c) sequences consisting of at least 20 contiguous  
 CC residues of (a); (d) sequences that hybridize to (a), under moderately  
 CC stringent conditions; (e) sequences having at least 75% or 90% identity  
 CC to (a); or (f) degenerate variants of (a). Polypeptides (ABP68596-  
 CC ABP68637) encoded by (I) and oligonucleotide can be used to detect cancer  
 CC in a patient and compositions comprising polypeptides, polynucleotides,  
 CC antibodies, fusion proteins, T cell populations and antigen presenting  
 CC cells expressing the polypeptide are useful in treating pancreatic cancer  
 CC and stimulating an immune response. The polynucleotides can be used as  
 CC probes or primers for nucleic acid hybridisation, in the design and  
 CC preparation of ribozyme molecules for inhibiting expression of the tumour  
 CC polypeptides and proteins in the tumour cells, in vaccines and for gene  
 CC therapy. Note: The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 2235 BP; 562 A; 560 C; 537 G; 576 T; 0 U; 0 Other;

# Alignment Scores:

Pred. No.: 5,59e-75 Length: 2235  
 Score: 1879.00 Matches: 370  
 Percent Similarity: 99.20% Conservative: 0  
 Best Local Similarity: 99.20% Mismatches: 3  
 Query Match: 98.38% Indels: 0  
 DB: 6 Gaps: 0

US-09-544-776-2 (1-373) x ABV94681 (1-2235)

QY 1 MetGluAspLeuAspGlnSerProLeuValSerSerAspSerProProArgProGln 20  
 DB 142 ATGGAGAGCTGGACGACGCTCTCTGGTCTCTCTGGACAGCCACCCCGCGCGAG 201  
 QY 21 ProAlaPheLysTyrglnPheValArgGluProGluAspGluGluGluGluGlu 40  
 DB 202 CCGCGTTCAAGTACCAGTTCGTGGAGGAGCCGAGGAGGAGGAGGAGGAGGAG 261  
 QY 41 GluGluGluAspGluAspGluAspGluGluGluGluGluGluGluGluGluGlu 60  
 DB 262 GAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 321  
 QY 61 AlaGlyLeuSerAlaAlaProValProThrAlaProAlaAlaGlyValaProLeuMetAsp 80  
 DB 322 GCCGGCTGTCCGGGGCCCGAGTGCACCGCCCTTCCGCGCGCGCGCGCGCGCGCG 381  
 QY 81 PheGlyAsnAspPheValProAlaProAlaProArgGlyPheLeuProAlaProVal 100  
 DB 382 TTCGGAATGACTTCGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGTC 441  
 QY 101 AlaProGluArgGlnProSerTrpAspProSerProValSerSerThrValProAlaPro 120  
 DB 442 GCCCGGAGCGGAGCGCGCTTGGGACCGAGCCGCGTGTCTGTCGACCGCGCGCCA 501  
 QY 121 SerPheLeuSerAlaAlaValSerProSerLysLeuProGluAspGluProPro 140  
 DB 502 TCCCGCTGTCTGCTGCGCGAGTCTCGCCCTCCAGCTCCCTGAGGAGGAGGAGGAGGAG 561  
 QY 141 AlaArgProProProProProProProProProProProProProProProProPro 160

DB 562 GCCCGGCTCCCTCCCTCCCGCGCAGCGGTGAGCCGCCCGCAGAGAGCCCGTGTGGACC 621  
 QY 161 ProProAlaProAlaProAlaProProSerThrProAlaAlaProLysArgArgGly 180  
 DB 622 CCGCAGGCGCGGCTCCCGCGCGCGCGCGCTCCACCGCGCGCGCGCGCGCGCGCG 681  
 QY 181 SerSerGlySerValValValValValValValValValValValValValValVal 200  
 DB 682 TCCTCGGCTCAGTGGTGTGACCTCTCTGTTACTGAGAGACATTAAGAACTGAGTG 741  
 QY 201 ValPheGlyAlaSerPheLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 220  
 DB 742 GTGTTGGTGGCAGCTATTCTCTGCTCTTCAATGACAGTATTGACGATTGTGAGCGTA 801  
 QY 221 ThrAlaTyrlleAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 240  
 DB 802 ACAGCTCATTTGGCTTGGCCCTCTCTGTGTGACCATCAGCTTTAGGATATACAAAGGT 861  
 QY 241 ValilleGlnAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 260  
 DB 862 GTGATCCAGCTATTCAGAAATCAGATGAGGCCACCCATTCAGGGCATATCTGGAATCT 921  
 QY 261 GluValAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 280  
 DB 922 GAAGTGTCTATATCTGAGGAGTGGTTCAGAAAGTACAGTAATCTGCTCTTGGTCATGTG 981  
 QY 281 AsnCysThrilleGlyGluLeuArgGluLeuValAspLeuValAspLeuValAspSerLeu 300  
 DB 982 AACTCCAGCATAAAGAACTCAGGCGCTCTCTTTAGTTGATGATTTAGTTAGTTCTCTG 1041  
 QY 301 LysPheAlaValLeuMetTrpValPheThrTyValGlyAlaAlaPheAsnGlyLeuThr 320  
 DB 1042 AGTTTGGAGTGTGATGTTGGTATTTACCTATGTTGGTGGCTTCTTTAATGCTGTACA 1101  
 QY 321 LeuLeuAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 340  
 DB 1102 CTACTGATTTGGCTCTCATTTCACTTCACTGTTCTCTGTTTATTTATGAACGGCATCAG 1161  
 QY 341 AlaGlnilleAspHisTyrlleuGlyLeuAlaAlaAlaAlaAlaAlaAlaAlaAla 360  
 DB 1162 GCACAGATAGATCATTTATCTAGGACTTGCATAAAGAATGTTAAAGATGCTATGCTAAA 1221  
 QY 361 lleGlnAlaLysilleProGlyLeuLysArgLysAlaGlu 373  
 DB 1222 ATCCAGCAAAATCCCTGGATTGAAGCGCAAGCTGHA 1260  
 RESULT 6  
 AAC64406  
 ID AAC64406 standard; cDNA; 2240 BP.  
 XX AAC64406;  
 AC AAC64406;  
 XX  
 DT 08-FEB-2001 (first entry)  
 XX  
 DE Human Nogo B nucleotide sequence SEQ ID NO:1.  
 XX  
 KW Human; Nogo B; cell stress response; hyperphosphorylated; brain tumour;  
 KW stress-phosphorylated endoplasmic reticulum protein; cytostatic;  
 KW gene therapy; cell growth; cellular stress response; neuron growth;  
 KW regulator of oxidative stress; inhibitor of neurite outgrowth;  
 KW axon regeneration; diagnosis; cancer; identification; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO200060083-A1.  
 PN  
 XX  
 PD 12-OCT-2000.  
 XX  
 XX 07-APR-2000; 2000WO-US009383.  
 PF  
 XX  
 PR 08-APR-1999; 99US-0128372P.  
 PR 21-JUN-1999; 99US-0140331P.  
 XX





PI Koenig-Hoffman K, Kazinski M, Schaefer R, Kesper B;  
XX WPI; 2003-542134/51.  
XX New nucleic acids involved in apoptosis, useful for diagnosis and  
PT treatment of e.g. tumors and degenerative disease, also related proteins,  
PT antibodies and modulators.  
XX  
PS Claim 1b; SEQ ID NO 303; 517pp; German.  
XX  
CC This invention describes novel nucleic acid molecules that are associated  
CC with apoptosis and encode a polypeptide and are derived from a normalised  
CC gene library (embryonic or liver) or clone collections, and the extent of  
CC apoptosis measured by cell death detection assay or the CPRG assay  
CC (measuring loss of membrane integrity). The products of the invention  
CC have cytosolic, neuroprotective, immunosuppressive, antineumatic,  
CC antiarthritic, dermatological, antiinflammatory, hepatotropic, virucide,  
CC neurotropic, anticonvulsant, antiparkinsonian, vasotropic,  
CC cerebroprotective and antialcoholic activity and can be used for gene  
CC therapy. The polynucleotides also related vectors, hosts for their  
CC extracts), encoded polypeptide (or their receptors) and/or agents that  
CC inhibit their activity (including antisense sequences) are used for  
CC treatment or prevention of tumours, autoimmune or degenerative diseases  
CC and viral infections, specifically leukaemia, carcinoma, sarcoma,  
CC multiple sclerosis, rheumatoid arthritis, diabetes, lupus, or infection  
CC with hepatitis or influenza viruses, Alzheimer's, Huntington's or  
CC Parkinson's diseases, reperfusion injury, stroke and alcoholic liver  
CC disease. Detection of the polynucleotides and derived polypeptides can  
CC also be used for diagnosis of these diseases. This sequence encodes an  
CC apoptosis-associated protein described in the disclosure of the  
CC invention.  
XX  
SQ Sequence 1599 BP; 354 A; 452 C; 422 G; 371 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 6,23e-75 Length: 1599  
Score: 1875.00 Matches: 369  
Percent Similarity: 98.93% Conservative: 0  
Best Local Similarity: 98.93% Mismatches: 4  
Query Match: 98.17% Indels: 0  
DB: 10 Gaps: 0

US-09-544-776-2 (1-373) x ADI62860 (1-1599)

QY 1 MetGluAspLeuAspGlnSerProLeuValSerSerAspSerProProArgProGln 20  
DB 116 ATGGAAGACTGGACGAGTCTCTCTGGTCTGTCGAGACAGCCACCCGCGCGAG 175

QY 21 ProAlaPheLysTyrrGlnPheValArgGluProGluAspGluGluGluGlu 40  
DB 176 CCGCGCTTCAAGTACCAGTTCGTGAGGAGCCGAGGAGGAGGAGGAGGAG 235

QY 41 GluGluGluAspGluAspGluAspGluGluGluGluGluGluGluGluGlu 60  
DB 236 GAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 295

QY 61 AlaGlyLeuSerAlaAlaProValProThrAlaProAlaAlaGlyAlaProLeuMetAsp 80  
DB 296 GCGCGGCTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 355

QY 81 PheGlyAsnAspPheValProAlaProAlaProArgGlyPheLeuProAlaProVal 100  
DB 356 TTCGGAATGACTTCTGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 415

QY 101 AlaProGluArgGlnProSerTrpAspProSerProValSerSerThrValProAlaPro 120  
DB 416 GCGCGGAGGCGGAGGCGGCTTGGGACCGGAGCCGCGTCTGTCGACCGCGCGCA 475

QY 121 SerPheLeuSerAlaAlaAlaValSerProSerLysLeuProGluAspGluProPro 140  
DB 476 TCCCCGCTGTCTGCTGCGGAGTCTCGCCCTCCAGCTCCGAGGAGGAGGAGGAG 535

QY 141 AlaArgProProProProProProAlaSerValSerProGlnAlaGluProValTrpThr 160

DB 536 GCCCGGCTCCCTCCCTCCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 595  
QY 161 ProProAlaProAlaProAlaProProSerThrProAlaAlaProLysArgGly 180  
DB 596 CGCGGAGGCGGCGGCTCCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 655  
QY 181 SerSerGlySerValValValAspLeuLeuTyrrTrpArgAspLysLysVal 200  
DB 656 TCCTCGGCTCAGTGCTGTTGACCTCTCTGACTCTGAGAGACATTAAAGAGCTGAGTG 715  
QY 201 ValPheGlyAlaSerLeuPheLeuLeuLeuSerLeuThrValPheSerLeValSerVal 220  
DB 716 GGTGTTGGTGGCAGGCTATTCTCTGCTGCTTTCATTGACAGATTTCAGCATTTGAGCGTA 775  
QY 221 ThrAlaTyrrLeuAlaLeuAlaLeuLeuSerValThrLysSerProArgLysLysGly 240  
DB 776 ACAGGCTCATTTGCTTGGCCCTGCTCTCTGTGACCATCAGCTTTAGGATATACAGGCT 835  
QY 241 ValLleGlnAlaLleGlnLysSerAspGluGlyHisProPheArgAlaTyrrLeuGluSer 260  
DB 836 GTGATCCAGCTATCCAGAAATCAGATGAGGCCACCCATTGAGGCATATCTGGAATCT 895  
QY 261 GluValAlaLleSerGluGluLeuValGlnLysTyrrSerAsnSerAlaLeuGlyHisVal 280  
DB 896 GAAGTTGCTATATCTGAGGAGTTGGTTTCAGAAAGTACAGTAATCTCTCTCTGTCATGTG 955  
QY 281 AsnCysThrLysGluLeuLeuArgLeuPheLeuValAspLeuValAspLeuValAspLeu 300  
DB 956 AACTGCAGATAAAGGAACTCAGAGGCGCTCTCTCTTAGTTGATGATTAGTTCTCTCTG 1015  
QY 301 LysPheAlaValLeuMetTrpValPheThrTyrrValGlyAlaLeuPheAsnGlyLeuThr 320  
DB 1016 AAGTTTGACGTGTTGATGCGGTATTTACCTATGTTGGTCCCTGTTTAAATGCTGACA 1075  
QY 321 LeuLeuLeuLeuAlaLeuLeuSerLeuPheSerValProValLysTyrrGluArgHisGln 340  
DB 1076 CTACTGATTTTGGCTCTCATTTCACTCTTCACTGTTCTCTGTTTATTTATGACGCGATCAG 1135  
QY 341 AlaGlnLleAspHisTyrrLeuGlyLeuAlaLeuLysAsnValLysAspAlaMetAlaLys 360  
DB 1136 GCACAGATAGATCATTTATCTAGGACTTGCAATAGAATGTTAAAGATGCTATGCTATAA 1195

QY 361 IleGlnAlaLysIleProGlyLeuLysArgLysAlaGlu 373  
DB 1196 ATCCAGCAAAATCCCTGGATTGAGCGCAAGCTGAA 1234

RESULT 8  
ID ABK90133 standard; DNA; 2052 BP.  
XX  
XX AC ABK90133;  
XX  
XX 21-OCT-2002 (first entry)  
XX  
XX DNA encoding human Nogob protein.  
XX  
XX Human; Nogob; BACE; acute neuronal injury; spinal injury; head injury;  
KW stroke; peripheral nerve damage; neoplastic disorder; glioblastoma;  
KW neuroblastoma; hyperproliferative disorder; dysproliferative disorder;  
KW cirrhosis; psoriasis; keloid formation; fibrocystic condition; cancer;  
KW tissue hypertrophy; central nervous system; axon regeneration; Nogob;  
KW Nogo-associated disease; metastasis; gene; ds.  
XX  
OS Homo sapiens.  
XX  
XX Key Location/Qualifiers  
FT 67..1188  
FT CDS /\*tag= a  
FT /product= "Human Nogob protein"  
XX  
XX WO200257483-A2.  
XX







PR 02-MAY-2000; 2000JP-00183865.  
 PR 07-JUL-2000; 2000EP-00114089.  
 XX (REAS-) RES ASSOC BIOTECHNOLOGY.  
 XX Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y,  
 PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;  
 XX WPI; 2004-204755/20.  
 DR P-PSDB; ADL31138.  
 XX  
 XX New oligonucleotide primers (830 cDNAs) useful for synthesizing full  
 PT length human cDNAs.  
 XX  
 XX Example 1; SEQ ID NO 3170; 1340pp; English.  
 XX  
 CC This invention relates to a novel primers useful for synthesizing full  
 CC length cDNA molecules that encode human proteins. Specifically, it refers  
 CC to secretory or membrane proteins that are potential therapeutic agents/  
 CC target molecules in the field of medicine, and in particular genes  
 CC encoding proteins that are associated with signal transduction,  
 CC glycoproteins and transcription. The present invention describes a method  
 CC for efficiently cloning a full length human cDNA from both the 5' and 3'  
 CC ends using the oligo-capping method. This polynucleotide sequence is a  
 CC full length human cDNA clone of the invention.  
 XX  
 SQ Sequence 1694 BP; 353 A; 484 C; 469 G; 388 T; 0 U; 0 Other;  
 XX

Alignment Scores:  
 Pred. No.: 2,48e-74 Length: 1694  
 Score: 1862.00 Matches: 369  
 Percent Similarity: 98.66% Conservations: 0  
 Best Local Similarity: 98.66% Mismatches: 4  
 Query Match: 97.49% Indels: 1  
 DB: 12 Gaps: 0

US-09-544-776-2 (1-373) x ADL31137 (1-1694)

QY 1 MetGluAspLeuAspGlnSerProLeuValSerSerSerAspSerProProArgProGln 20  
 DB 232 ATGGAGACCTGGACCACTCTCTCTGCTCTGCTCTGGACAGCCACCCCGGCCGAG 291  
 QY 21 ProAlaPheLysTyrGlnPheValArgGluProGluAspGluGluGluGluGlu 40  
 DB 292 CCGCGCTCAAGTACCACTGCTGAGGAGCCCGGAGGAGGAGGAGGAGGAGGAG 351  
 QY 41 GluGluAspGluAspGluAspGluGluGluGluGluGluGluGluGluGluGlu 60  
 DB 352 GAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 411  
 QY 61 AlaGlyLeuSerAlaAlaProValProThrAlaProAlaAlaGlyAlaProLeuMetAsp 80  
 DB 412 GCGGGGCTGTCGGGGCCCGAGTGGCCACCGCCCTGCGCCGGGGGGCCCTGTATGGAC 471  
 QY 81 PheGlyAsnAspPheValProProAlaProAlaProArgGlyPheLeuProAlaProVal 100  
 DB 472 TTCGAAATGACTTGGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGGT 531  
 QY 101 -AlaProGluArgGlnProSerTrpAspProSerProValSerSerThrValProAlaPr 120  
 DB 532 CGCCCCGGAGCGGAGCGGCTTGGGACCCGAGCCCGGTGTGTCGACCGGTGCGCGGCC 591  
 QY 120 oSerPheLeuSerAlaAlaValSerProSerLysLeuProGluAspAspGluProPr 140  
 DB 592 ATCCCGCTGTCTGCTGCCGAGTCTCGCCCTCCAGCTCCCTGAGGACGAGCGCTCC 651  
 QY 140 oAlaArgProProProProProAlaSerValSerProGlnAlaGluProValTrpTh 160  
 DB 652 GGCCCGGCTCCCTCTCCCGCGGCGAGCGTGGAGCCCGGAGGAGCGCGTGTGGAC 711  
 QY 160 rProProAlaProAlaProAlaProProSerThrProAlaProLysArgGlu 180  
 DB 712 CCGGCCAGCCCGGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGGG 771

QY 180 YSerSerGlySerValValValAspLeuLeuTyrTrpArgAspIleLysLysThrGlyVa 200  
 DB 772 CTCTCGGGCTCAGTGGTGTGTGACCTCTGTACTGGAGAGACATTAAGAAGACTGGAGT 831  
 QY 200 lValPheGlyAlaSerLeuPheLeuLeuSerLeuThrValPheSerIleValSerVa 220  
 DB 832 GGTGTTGGTGTGAGCTATTCCTGCTGCTTTCATTGACAGTATTTCAGCATTTGAGCGT 891  
 QY 220 lThrAlaTyrIleAlaLeuAlaLeuSerValThrIleSerProArgIleTyrLysG 240  
 DB 892 AACAGCTCATTCCTGCTGCTCTCTGTGACCATCAGCTTAGGATATACAAGG 951  
 QY 240 YValIleGlnAlaIleGlnLysSerAspGluGlyHisProPheArgAlaTyrLeuGluSe 260  
 DB 952 TGTGATCCAGCTATCCAGAAATCAGATGAAGCCACCCATTTCAGGGCATATCTGGAATC 1011  
 QY 260 rGluValAlaIleSerGluGluLeuValGlnLysTyrSerAsnSerAlaLeuGlyHisVa 280  
 DB 1012 TGAAGTTGCTATATCTCAGGAGTGGTTCAGAAAGTACAGTAATTCCTGCTCATGT 1071  
 QY 280 lAsnCysThrIleLysGluLeuArgGluPheLeuValAspAspLeuValAspSerLe 300  
 DB 1072 GAATCGACGATTAAGAACTCAGGCGCTCTTCTTAGTTGATGATTTAGTTGATTTCTCT 1131  
 QY 300 uLysPheAlaValLeuMetTrpValPheThrTyrValGlyValAlaLeuPheGlnGlyLeuTh 320  
 DB 1132 GAAGTTTGCAGTGTGATGGGTATTTACCTATGTTGGTGCCTTGTTTAATGCTGTAC 1191  
 QY 320 rLeuLeuIleLeuAlaLeuIleSerLeuPheSerValProValIleTyrGluArgHisG 340  
 DB 1192 ACTACTGATTTGGCTCTCATTTCACTTTCAGTGTTCCTGTTTATTTATGAACGGCATCA 1251  
 QY 340 nAlaGlnIleAspHisTyrLeuGlyLeuAlaAsnLysAsnValLysAspAlaMetAlaLy 360  
 DB 1252 GGCACAGATAGATTAICTAGGACTTGCATAATAGAATGTTAAGATGCTATGGCTAA 1311  
 QY 360 sIleGlnAlaLysIleProGlyLeuLysArgLysAlaGlu 373  
 DB 1312 AATCCAAAGCAAAATCCCTGGATTGAAGCGCAAGCTGAA 1351

RESULT 11  
 AD08386  
 ID AAD08386 standard; cDNA; 1683 BP.  
 XX  
 AC AAD08386;  
 DT 09-AUG-2001 (first entry)  
 XX  
 DE Human secreted protein-encoding gene 42 cDNA clone HAGFT48, SEQ ID NO:52.  
 DE  
 KW Human; secreted protein; proliferative disorder; cancer; tumour;  
 KW foetal abnormality; developmental abnormality; haematopoietic disorder;  
 KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;  
 KW inflammation; allergy; neurological disorder; Alzheimer's disease;  
 KW Parkinson's disease; cognitive disorder; schizophrenia; asthma;  
 KW skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;  
 KW cardiovascular disorder; angioneurotic disorder; kidney disorder;  
 KW gastrointestinal disorder; pregnancy-related disorder;  
 KW endocrine disorder; infection; wound healing; vulnerability; cell culture;  
 KW chemotaxis; food additive; gene therapy; binding partner identification;  
 KW ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 CDS 830..1192  
 FT /tag= a  
 FT /product= "Human secreted protein precursor"  
 FT /note= "CDS does not include start codon"  
 FT /partial  
 FT sig\_peptide  
 FT 830..898  
 FT /\*tag= b

FT mat\_peptide 899..1189  
 PT /\*tag= c  
 FT /product= "Mature human secreted protein"  
 XX  
 PN WO200077022-A1.  
 XX  
 PD 21-DEC-2000.  
 XX  
 XX 01-JUN-2000; 2000WO-US015136.  
 PF 11-JUN-1999; 99US-0138629P.  
 PR (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PA Rosen CA, Ruben SM, Komatsoulis GA;  
 PI WPI; 2001-367020/38.  
 XX P-PsDB; AAE03939.  
 DR  
 DR  
 XX  
 XX Nucleic acids encoding 50 human secreted polypeptides, useful for  
 PT preventing, diagnosing and/or treating diseases, e.g. Parkinson's  
 PT disease, botulism, cancers and Schmitar syndrome.  
 FT  
 PS Claim 1; Page 520; 614pp; English.  
 XX  
 CC AAD08345-AAD08394 represent cDNAs corresponding to 50 human secreted  
 CC protein genes and AAE03898-AAE03947 represent the proteins they encode.  
 CC AAE03948-AAE03996 represent human secreted protein fragments or variants.  
 CC The genes and their secreted proteins are useful for preventing, treating  
 CC or ameliorating medical conditions, e.g., by protein or gene therapy.  
 CC Pathological conditions can be diagnosed by determining the amount of the  
 CC new protein in a sample or by determining the presence of mutations in  
 CC the new genes. Specific uses are described for each of the 50 genes,  
 CC based on the tissues in which they are most highly expressed, and include  
 CC developing products for the diagnosis or treatment of proliferative  
 CC disorders, cancer, tumours, foetal and developmental abnormalities,  
 CC haematopoietic disorders, diseases of the immune system, AIDS, autoimmune  
 CC diseases (e.g. rheumatoid arthritis), inflammation, allergies, disease),  
 CC neurological disorders (e.g., Alzheimer's disease, Parkinson's disease),  
 CC cognitive disorders, schizophrenia, asthma, skin disorders (e.g.,  
 CC psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders,  
 CC angioecnic disorders, kidney disorders, gastrointestinal disorders,  
 CC pregnancy-related disorders, endocrine disorders, and infectious. The  
 CC proteins can also be used to aid wound healing and epithelial cell  
 CC proliferation, to prevent skin aging due to sunburn, to maintain organs  
 CC before transplantation, for supporting cell culture of primary tissues,  
 CC to regenerate tissues, to identify their cognate ligands or binding  
 CC partners, and in chemotaxis, and can be used as a food additive or  
 CC preservative to modify storage properties. Antibodies specific for a  
 CC protein of the invention can be used in alleviating symptoms associated  
 CC with the disorders mentioned above, and in diagnostic immunoassays e.g.,  
 CC radioimmunoassay or enzyme linked immunosorbent assay (ELISA). The  
 CC present sequence represents a human secreted protein-encoding cDNA of the  
 CC invention  
 XX  
 SQ Sequence 1683 BP; 369 A; 489 C; 464 G; 358 T; 0 U; 3 Other;  
 Alignment Scores:  
 Pred. No.: 3,45e-68 Length: 1683  
 Score: 1723.50 Matches: 346  
 Percent Similarity: 92.76% Conservative: 0  
 Best Local Similarity: 92.76% Mismatches: 4  
 Query Match: 90.24% Indels: 24  
 DB: 4 Gaps: 1  
 US-09-544-776-2 (1-373) x AAD08386 (1-1683)  
 QY 1 MetGluAspLeuAspGlnSerProLeuValSerSerSerAspSerProArgProGln 20  
 DB 233 ATGGAGAGCTGGACCAAGTCTCTCTGGTCTCTGGACAGCCACCCCGCCGCGAG 292  
 QY 21 ProAlaPheLysTyrGlnPheValArgGluProGluAspGluGluGluGlu 40  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 293 CCCGGCTTCAAGTACCAGTTCTGAGGGAGCCCGAGGACGAGGAGGAGGAG 352  
 QY 41 GluGluGluAspGluAspGluAspLeuGluGluValLeuGluArgLysProAla 60  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 353 GAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 412  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 61 AlaGlyLeuSerAlaAlaProValProThrAlaProAlaAlaGlyAlaProLeuMetAsp 80  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 413 GCGGGGCTGTCCGGGCCCCAGTGCACCGGCCCTCCCGCGCGCGCCCTCATGGAC 472  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 81 PheGlyAsnAspPheValProProAlaProArgGlyPheLeuProAlaAlaProProVal 100  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 473 TTTCGGAATGACTTCGTGCGCGCGCGCCCGGGGACCCCTGCGGCGCTCCCGCGTC 532  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 101 AlaProGluArgGlnProSerTrpAspProSerProValSerSerTrpValProAlaPro 120  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 533 GCGCGGAGCGGCGGCGCTTGGGACCCGAGCCCGGTGCTGCGACCGTCCCGCGCCA 592  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 121 SerPheLeuSerAlaAlaValSerProSerLysLeuProGluAspAspGluProPro 140  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 593 TCCCGCTGTCTGCTGCGCGAGTCTCGCCCTCCAGCTCCCTGAGGACGAGGAGCTCCG 652  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 141 AlaArgProProProProProAlaSerValSerProGlnAlaGluProValTrpThr 160  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 653 GCGCGGCTCCCGCTCTCTCCCGCGCGAGCGGTGAGCCCGCAGGAGCCCGTGTGACC 712  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 161 ProProAlaProAlaProAlaProProSerTrpProAlaAlaProLysArgArgGly 180  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 713 CCGCGAGCGCGCGCTCCCGCGCGCCCTCCACCGCGCGCGCCCGCGAGGAGGAG 772  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 181 SerSerGlySerValValValAspLeuLeuTrpArgAspLysLysLysThrGlyVal 200  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 773 TCTCTGGGCTCAGTGGTGTGTGACCTCTCTGACTGAGAGACATTAAGAGAGCTGAGTG 832  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 201 ValPheGlyAlaSerLeuPheLeuLeuSerLeuThrValPheSerLysValSerVal 220  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 833 GTGTTTGGTGCAGGCTATTCTCTGCTCTTTCATTCAGACAGTATTCAGCATTTGAGCGTA 892  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 221 ThrAlaTyrIleAlaLeuAlaLeuSerValThrIleSerProArgLysLysGly 240  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 893 ACAGGCTACATTCGCTTGGCCCTGCTCTCTGACCATCAGCTTTAGGATATACAGGGT 952  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 241 ValIleGlnAlaIleGlnLysSerAspGluGlyHisProPheArgAlaTyrLeuGluSer 260  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 953 GTGATCCAGCTATCCAGAAATCAGATGAAGCCGCCACCCATTGAGGATATCTGGAATCT 1012  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 261 GluValAlaIleSerGluGluValGlnLysTyrSerAsnSerAlaLeuGlyHisVal 280  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 1013 GAAGTTGCTATATCTGAGGAGTTGGTTCAGAAAGTACAGTAATCTGCTCTGTCATGTG 1072  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 281 AsnCysThrIleLysGluLeuArgLeuPheLeuValAspLeuValAspSerLeu 300  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 1073 AACTGCAGATAAAGAACTCAGGCGCTCTCTCTTAGTTGATGATTAGTTGATCTCTG 1132  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 301 LysPheAlaValLeuMetTrpValPheThrTyrValGlyAlaLeuPheAsnGlyLeuThr 320  
 |||||  
 Db 1133 AGCT----- 1137  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 321 LeuLeuIleLeuAlaLeuIleSerLeuPheSerValProValIleTyrGluArgHisGln 340  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 1138 -----CTCATTTTCACTTTCAGTGTTCCTGTTTATTTATGAACGCGATCAG 1182  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 341 AlaGlnIleAspHisTyrLeuGlyLeuAlaAsnLysAsnValLysAspAlaMetAlaLys 360  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 1183 GCACAGATAGATCATTATCTAGGACTTGCAATAGAAATGTTAAGATGCTATGGCTAA 1242  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 361 IleGlnAlaLysIleProGlyLeuLysArgLysAlaGlu 373  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 1243 ATCCAGCAAAATCCCTGGATTGAAGCGCAAGCTGAA 1281  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 RESULT 12  
 ADB85284  
 ID ADB85284 standard; DNA; 2782 BP.



XX AAZ56886;  
AC  
XX  
DT 25-APR-2000 (first entry)  
XX  
DE Human MAGI polypeptide encoding DNA.  
XX  
XX MAGI protein; neuroendocrine-specific protein; neuropathy; human;  
KW spinal injury; neuronal degeneration; neuromuscular disorder; cancer;  
KW psychiatric disorder; developmental disorder; inflammatory disorder;  
KW stroke; cytostatic; cerebroprotective; neuroprotective; ds.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..3579  
FT /tag= a  
FT /product= "MAGI polypeptide"  
XX  
XX WO200005364-A1.  
PD 03-FEB-2000.  
XX  
XX 21-JUL-1999; 99WO-GB002360.  
XX  
XX 22-JUL-1998; 98GB-00016024.  
XX 19-JUL-1999; 99GB-00016898.  
XX  
XX (SMIK ) SMITHKLINE BEECHAM PLC.  
XX  
XX Michalovich D, Prinjha RK;  
XX WPI; 2000-182693/16.  
DR P-PSDB; AAY56967.  
XX  
XX Novel polypeptides related to neuroendocrine-specific proteins and  
PT polynucleotides useful for diagnosis of various diseases and for  
PT treatment of cancer and neurological disorders.  
XX  
PS Claim 5; Page 19-20; 35pp; English.  
XX  
CC The invention relates to human MAGI protein, which is similar to  
CC neuroendocrine-specific protein. The MAGI protein can be expressed by  
CC standard recombinant methodology. The MAGI polypeptides, polynucleotides  
CC and antibodies are useful for treating diseases, including neuropathies,  
CC spinal injury, neuronal degeneration, neuromuscular disorders,  
CC psychiatric disorders and developmental disorders, cancer, stroke and  
CC inflammatory disorders. The polynucleotide is also useful for chromosome  
CC localization and for tissue expression studies. The present sequence  
CC represents a DNA encoding the human MAGI protein  
XX  
SQ Sequence 3579 BP; 1074 A; 803 C; 812 G; 890 T; 0 U; 0 Other;  
  
Alignment Scores:  
Pred. No.: 3 53e-56 Length: 3579  
Score: 1459.50 Matches: 370  
Percent Similarity: 31.04% Conservative: 0  
Best Local Similarity: 31.04% Mismatches: 3  
Query Match: 76.41% Indels: 819  
DB: 3 Gaps: 1  
  
US-09-544-776-2 (1-373) x AAZ56886 (1-3579)  
  
QY 1 MetGluAspLeuAspGlnSerProLeuValSerSerSerAspSerProProArgProGln 20  
Db 1 ATGGAAGACCTGGACCAAGTCTCTCTGGTCTCGTCTCGACAGCCACCCGGCGCAG 60  
  
QY 21 ProAlaPheIysTyrGlnPheValArgGluProGluAspGluGluGluGlu 40  
Db 61 CCCGCGTTTCAATACAGTTCGTGAGGAGCCCGCAGCAGGAGGAGGAGGAGGAG 120  
  
QY 41 GluGluAspGluAspGluAspGluGluGluGluGluGluGluGluGluGluGlu 60  
|||||

Db 121 GAAGAGGAGGACGAGGACGAAGACCTGGAGGAGCTGGAGGTCTGGAGGAGGAGCCCGCC 180  
QY 61 AlaGlyLeuSerAlaAlaProValProThrAlaProAlaAlaAlaGlyAlaProLeuMetAsp 80  
Db 181 GCCGGGCTGTCCGGCGCCCGCCAGTGCACCGCCCTGCCCGCGCGCCCTGTATGGAC 240  
QY 81 PheGlyAsnAspPheValProProAlaProAlaProAlaGlyPheLeuProAlaAlaProVal 100  
Db 241 TTCGGAATAGACTTGTGCCCGCGCGCCCGCGGAGACCCCTGCCCGCGCTCCCCCGTC 300  
QY 101 AlaProGluArgGlnProSerTrpAspProSerProValSerSerThrValProAlaPro 120  
Db 301 GCCCGGAGCGGACGCGCTTTGGGACCCGACCGCGGTGCTGCGACCGTGCCTGCCGCCA 360  
QY 121 SerPheLeuSerAlaAlaAlaValSerProSerIysLeuProGluAspGluProPro 140  
Db 361 TCCCGCTGTCTGCTGCCGACGTCTCGCCCTCCAAGCTCCCTGAGGACACGAGCCTCG 420  
QY 141 AlaArgProProProProProProAlaSerValSerProGlnAlaGluProValTrpThr 160  
Db 421 GCCCGGCTTCCCCCTCTCCCCCGGCGGAGGAGCCCGCAGGAGCCCGTGTGGACC 480  
QY 161 ProProAlaProAlaAlaProProSerThrProAlaAlaProIysArgGly 180  
Db 481 CGCCAGCCCGGCTCCCGCGCGCCCTCCACCCCGCGCGCCCAAGCGCAGGGGC 540  
QY 181 SerSerGlySerVal----- 185  
Db 541 TCCTCGGCTCAGTGGATGAGACCCCTTTTGTCTTCTGCTGCTGCTGAGCCTGTGATA 600  
QY 185 ----- 185  
Db 601 CGCTCTCTGCAGAAATATGGACTTGAAGGAGCAGCAGGTAAACACTATTTCCGGTGT 660  
QY 185 ----- 185  
Db 661 CAAGAGGATTTCCCATCTGCTGCTTGAACTGCTGCTTCTCTTCTTCTCTCTCTCTCT 720  
QY 185 ----- 185  
Db 721 CTCTCAGCGCTTTCTTCAAAGAAATGAATACCTTGGTAAATTTGTCAACAGTATTACCC 780  
QY 185 ----- 185  
Db 781 ACTGAAGNACACTTCAAGAAATGTCAGTGAAGCTTCTAAAGAGTCTCAGAGAAGGCA 840  
QY 185 ----- 185  
Db 841 AAAACTCTACTCATAGATAGAGATTAAACAGAGTTTTCAGAAATTAGAAATCTCAGAAATG 900  
QY 185 ----- 185  
Db 901 GGATCATCGTTCAGTGTCTCTCCAAAGCAGAAATCTGCCGTAAATAGTAGCAAAATCTTAG 960  
QY 185 ----- 185  
Db 961 GAAGAAATATCGTGAAAAATAAAGATGAAGAGAGAGTGTAGTAGTAATACATCCTT 1020  
QY 185 ----- 185  
Db 1021 CATAATCAACAGAGTTACCTACAGCTCTTACTAAATTTGGTTAAAGAGGATGAAGTTGTG 1080  
QY 185 ----- 185  
Db 1081 TCTTCAGAAAAAGCAAAAGACAGTTTTAAATGAAAAGAGAGTTGCAGTGAAGCTCCTATG 1140  
QY 185 ----- 185  
Db 1141 AGGAGGAATATGCAGACTTCAACACCTTTGAGCGAGTATGGAAAGTGAAGATAGTAAG 1200  
QY 185 ----- 185  
Db 1201 GAAGATAGTATATGTTGGTCTGCGAGTAAATTCGAGAGCAACTTGGAAAGATAAGTG 1260

QY 185 ----- 185  
Db 1261 GATAAAAAATGTTTTCAGATAGCCTTGAGCAAACTAATCACGAAAAAGATAGTGAGAGT 1320  
QY 185 ----- 185  
Db 1321 AGTAATGATGATACTTCTTTCCAGTACGCCAGAGAGGTATAAAGGATCGTCCAGGAGCA 1380  
QY 185 ----- 185  
Db 1381 TATATCATGTGCTCCCTTTAACCCAGCAGCACTGAGAGCATTGCAACAAACATTTTT 1440  
QY 185 ----- 185  
Db 1441 CCTTTGTTAGGATCCTACTTCAGAAAAAAGACCCGATGAAAAAATAAGAGAAAG 1500  
QY 185 ----- 185  
Db 1501 AAGGCCAAATAGTAACAGAGAAGAAATACTAGCACCAAAACATCAAAACCCTTTTCTTGTA 1560  
QY 185 ----- 185  
Db 1561 GCAGCAGAGATTCTGAGACAGATTATGTCAACACAGATAAATTAAACAAGGTGACTGAG 1620  
QY 185 ----- 185  
Db 1621 GAAGTCGTGCAAAACATGCTGAAGGCTGACTCCAGATTAGTACAGGAAGCATGTGAA 1680  
QY 185 ----- 185  
Db 1681 AGTGAATTGAATGAAGTTACTGGTACAAAGATTGCTTATGAAACAAAAATGGACTTGGTT 1740  
QY 185 ----- 185  
Db 1741 CAAACATCAGAAGTTATGCAAGAGTCACCTATCTCTGCACGACAGCTTTGGCCCATCATTT 1800  
QY 185 ----- 185  
Db 1801 GAAGAGTCAGAAGCTACTCCTTCACCAGTTTTCGCTGACATTGTTATGGAAGCACCATTG 1860  
QY 185 ----- 185  
Db 1861 AATTCTGCAGTTCTTAGTGCTGGTCTTCGGTGATACAGCCAGCTCATCACCATTAGAA 1920  
QY 185 ----- 185  
Db 1921 GCTTCTTCAGTTAATTATGAAGCATAAACATGAGCCTGMAAAACCCCCACCATATGAA 1980  
QY 185 ----- 185  
Db 1981 GAGGCCATGAGTGTATCACTAAAAAAGTATCAGGAATAAGGAAGAAATTAAGAGCCT 2040  
QY 185 ----- 185  
Db 2041 GAAAAATTAATGAGCTCTTCAAGAAACAGAAGCTCCTTATATATCTATTGATGTGAT 2100  
QY 185 ----- 185  
Db 2101 TTAATTAAAGAAACAAAGCTTCTGTGTAACAGCTCCGAGTTCTCTGATTATTCAGAA 2160  
QY 185 ----- 185  
Db 2161 ATGCCAAAGTTGAACAGCCAGTCGCTGATCATTTCTGAGCTAGTTGAAGATTCTCACCT 2220  
QY 185 ----- 185  
Db 2221 GATTCTGAACAGTTGACTTATTTAGTGATGATTCAATACCTGACCGTTCCACAAAAACAA 2280  
QY 185 ----- 185  
Db 2281 GATGAAACTGTGATGCTTGTGAAAGAAAGTCTCACTGAGACTTCATTTGAGTCAATGATA 2340

QY 185 ----- 185  
Db 2341 GAATATGAAAAATAAGAAAAAACTCAGTGTCTTCCACCTGAGGAGGAAAGCCATATTTG 2400  
QY 185 ----- 185  
Db 2401 GAATCTTTTAAGCTCAGTTTAGATAACAAAAAGATACCTGTGTACTGTATGAAGTTTCA 2460  
QY 185 ----- 185  
Db 2461 ACATTGAGCAAAAAGAGAAAAATTCCTTTGCAGATGGAGGAGCTCAGTACTGCAGTTTAT 2520  
QY 185 ----- 185  
Db 2521 TCAAATGATGACTTATTTATTTCTAAGGAAGCACAGATAAGAGAACTGAAACGTTTCA 2580  
QY 185 ----- 185  
Db 2581 GATTCAATCTCCAATTGAAATTTATAGATGAGTTCCCTACATTTGATCAGTTCTTAAACATGAT 2640  
QY 185 ----- 185  
Db 2641 TCATTTTCTAAATTAGCCAGGGAATATACTGACCTAGAAAGTATCCACAAAAAGTGAAATT 2700  
QY 185 ----- 185  
Db 2701 GCTAATGCCCCGATGGAGCTGGGTCAITTCCTTTGCACAGAAATGGCCCATGACCTTTCT 2760  
QY 185 ----- 185  
Db 2761 TTGAAGAACATACAAACCAAGTTGAAGAGAAAAATCAGTTTCTCAGATGACTTTCTTAA 2820  
QY 185 ----- 185  
Db 2821 AATGGGTCTGCTACATCAAAAGGTGCTCTTATTGCTCCAGATGTTTCTGCTTTGGCCACT 2880  
QY 185 ----- 185  
Db 2881 CAAGCAGATAGAGAGCATAGTTAAACCCAAAGTTCTTGTGAAGAAGCTGAGAAAAAA 2940  
QY 185 ----- 185  
Db 2941 CTTCTTCCGATACAGAAAAAGAGGACAGATCACCATCTGCTATATTTTTCAGCAGAGCTG 3000  
QY 186 ----- ValValAspLeuLeuTyrTrpArgAspIleValSerGlyValVal 201  
3001 AGTAAAACTTCAGTTGTTGACCTCTGTAAGAGACATTAAGAGACTGAGTGGTG 3060  
QY 202 PheGlyAlaSerLeuPheLeuSerLeuThrValPheSerIleValSerValThr 221  
3061 TTTGGTGCAGGCTATTCTCTGCTTTCATTGACAGTATTCAGCATTTGTGAGCGTAACA 3120  
QY 222 AlaTyrIleAlaLeuAlaLeuSerValThrIleSerProArgIleTyrIleVal 241  
3121 GCTCATATGCTTGGCCCTGCTCTCTGTGACCATCAGCTTTAGGATATACAAAGGGTGTG 3180  
QY 242 IleGlnAlaIleGlnIleValSerGlyHisProPheArgAlaTyrLeuGluSerGlu 261  
3181 ATCCAGCTATCCAGAAATCAGATGAAGGCCACCCATTGAGGCATATCTGGAATCTGAA 3240  
QY 262 ValAlaIleSerGluIleValGlnIleValGlnIleValGlnIleValGlnIleValGln 281  
3241 GTTGTATATCTGAGGAGTTGGTTTCAGAAAGTACAGTAATTCCTGCTTGTGTCATGTGAAC 3300  
QY 282 CysThrIleValGluLeuArgLeuPheLeuValAspLeuValAspSerLeuVal 301  
3301 TGCACGATAAAGAACTCAGGCGCTCTTCTTAGTTGATGATTAGTTGATTCCTGGAAG 3360  
QY 302 PheAlaValLeuMetTrpValPheThrTyrValGlyAlaLeuPheAsnGlyLeuThrLeu 321  
3361 TTTGCAAGTTGATGTGGGTATTTACCTATGTTGGTCCCTGTTTAAATGGTCTGACACTA 3420  
QY 322 LeuIleLeuAlaLeuIleSerLeuPheSerValProValIleTyrGluArgHisGlnAla 341



Db 3421 CTGATTTGGCTCTCATTTTCACTCTTTCAGTGTCTCTGTTATTTATGAACGGCATCAGGCG 3480  
QY 342 GlnIleAspHisTyrLeuGlyLeuAlaAsnLysAsnValLysAspAlaMetAlaLysIle 361  
Db 3481 CAGATAGATCATTTATCTAGGACTTGCATAATAAGAAATGTTAAAGATGCTATGGCTAAATC 3540  
QY 362 GlnAlaLysIleProGlyLeuLysArgLysAlaGlu 373  
Db 3541 CAAGCAAAATTCCTGGATTGAAGCCGAAGCTGAA 3576  
RESULT 14  
AAF90324  
ID AAF90324 standard; cDNA; 3579 BP.  
XX AC AAF90324;  
DT 23-JUL-2001 (first entry)  
XX DE Human NOGO-A cDNA.  
XX KW NOGO-A; human; chromosome 2p21; neuropathy; spinal injury; brain injury; stroke; neuronal degeneration; Alzheimer's disease; Parkinson's disease;  
KW neuromuscular disorder; psychiatric disorder; developmental disorder;  
KW neuroprotective; nootropic; neuroleptic; antiparkinsonian;  
KW cerebroprotective; neuroleptic; diagnosis; therapy; ss.  
XX OS Homo sapiens.  
XX PN WO200136631-A1.  
XX PD 25-MAY-2001.  
XX PF 14-NOV-2000; 2000WO-GB004345.  
XX PR 15-NOV-1999; 99GB-00026995.  
XX PR 24-JAN-2000; 2000GB-00001550.  
XX PA (SMIK ) SMITHKLINE BEECHAM PLC.  
XX PI Michalovich D, Prinjha R;  
XX DR WPI; 2001-343822/36.  
XX DR P-PSDB; AAB82349.  
XX PT New polypeptide designated NOGO-C is a splice variant of the human NOGO  
PT gene and may be useful in the treatment of neural disorders including  
PT Alzheimer's and Parkinson's diseases.  
XX PS Disclosure; Page 25-26; 25pp; English.  
XX CC The present sequence is that of cDNA encoding human NOGO-A (see  
CC AAB82349). NOGO-A is a previously known splice variant of the human NOGO  
CC gene on chromosome 2p21. NOGO-A cDNA was obtained by PCR amplification of  
CC human spinal cord cDNA. The invention relates to a novel splice variant,  
CC NOGO-C (see AAF90323). It provides NOGO-C polypeptides and  
CC polynucleotides, and methods for producing such polypeptides by  
CC recombinant techniques. Also disclosed are methods for utilising NOGO-C  
CC polypeptides and polynucleotides in the treatment of diseases including  
CC neuropathies, spinal injury, brain injury, stroke, neuronal degeneration,  
CC for example Alzheimer's disease and Parkinson's disease, neuromuscular  
CC disorders, psychiatric disorders and developmental disorders. Also  
CC provided are methods for identifying agonists and agonists for use in  
CC treating conditions associated with NOGO-C imbalance, and diagnostic  
CC assays for detecting diseases associated with inappropriate NOGO-C  
CC activity or levels  
SQ Sequence 3579 BP; 1074 A; 803 C; 812 G; 890 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 353e-56 Length: 3579  
Score: 1459.50 Matches: 370  
Percent Similarity: 31.04% Conservative: 0

Best Local Similarity: 31.04% Mismatches: 3  
Query Match: 76.41% Indels: 819  
DB: 4 Gaps: 1  
US-09-544-776-2 (1-373) x AAF90324 (1-3579)  
QY 1 MetGluAspLeuAspGlnSerProLeuValSerSerSerSerSerProArgProGln 20  
Db 1 ATGGAAGACCTGGACCACTCTCTGGTCTCGTCTCGGACAGCCACCCCGGCGCAG 60  
QY 21 ProAlaPheLysTyrGlnPheValArgGluProGluAspGluGluGluGluGlu 40  
Db 61 CCGCGTTCAAGTACCAGTTCTGTGAGGAGCCGAGGACGAGGAGGAGGAG 120  
QY 41 GluGluGluAspGluAspGluAspLeuGluGluGluGluGluGluGluGluGlu 60  
Db 121 GAAGAGGAGGACGAGGACGAGACCTGGAGGAGCTGGAGGTCTGGAGAGGAGCCGCC 180  
QY 61 AlaGlyLeuSerAlaAlaProValProThrAlaProAlaAlaGlyAlaProLeuMetAsp 80  
Db 181 GCCGGGCTGTCCGCGCCCGCCAGTGCACCCCTGCCGCGCGCGCCCTGATGGAC 240  
QY 81 PheGlyAsnAspPheValProProAlaProAlaProArgGlyPheLeuProAlaAlaProVal 100  
Db 241 TTCGGAATGACTTCTGTCGCGCGCGCGCGCGCGGACCCCTGCCGCGCTCCCCCGCTC 300  
QY 101 AlaProGluArgGlnProSerTrpAspProSerProValSerSerThrValProAlaPro 120  
Db 301 GCCCGGAGCGGACGCGCTTGGGACCGGAGCCGCTGCTCGACCGTGGCCCGGCCA 360  
QY 121 SerPheLeuSerAlaAlaAlaValSerProSerLysLeuProGluAspAspGluProPro 140  
Db 361 TCCCGCTGTCTGCTGCGCGCAGTCTGCGCTCCAGCTCCCTGAGGACGACGAGCTCCG 420  
QY 141 AlaArgProProProProProProAlaSerValSerProGlnAlaGluProValTrpThr 160  
Db 421 GCCCGCTCTCCCTCTCTCCCGCGCAGCGTGAGCCCGCGGAGGAGCCGTGTGGACC 480  
QY 161 ProProAlaProAlaProAlaProProSerThrProAlaAlaProLysArgGly 180  
Db 481 CCGCAGCCCGCGCTCCCGCGCGCTCCCTCCACCGCGCGCGCCCAAGCGCAGGGGC 540  
QY 181 SerGlySerVal----- 185  
Db 541 TCTCGGCTCAGTGGATGAGACCCCTTTTGTCTCTCTGCTGCTGCTGAGCCTGTGATA 600  
QY 185 ----- 185  
Db 601 CGCTCTCTGCAGAAAATATGGACTTGAAGGAGCAGCCAGGTAACTATTATTTCCGCTGGT 660  
QY 185 ----- 185  
Db 661 CAAGAGGATTTCCCATCTGCTCTGTTGAACTGCTGCTTCTTCTTCTCTGCTCTCTCT 720  
QY 185 ----- 185  
Db 721 CTCTCAGCGCTTCTTCAAGAACATGATACCTTGGTAATTTGTCAACAGTATTACCC 780  
QY 185 ----- 185  
Db 781 ACTGAAGGAACACTTCAAGAAAATGTCAAGAACTTCTAAAGAGGTCTCAGAGAAGGCA 840  
QY 185 ----- 185  
Db 841 AAACTCTACTCATAGATAGATTTAACAGAGTTTTCAGAAATTAGATCTCAGAAATG 900  
QY 185 ----- 185  
Db 901 GGATCATGTTGAGTGTCTCTCAAAAGCAGAACTCGCGGTAATAGTAGCAATCTTAGG 960  
QY 185 ----- 185  
Db 961 GAAGAAATAATCGTGAAAAATAAAGATGAAGAGAAGTTAGTTAGTAGTAATAACATCCTT 1020





Db 541 TCCTGGGCTCAGTGGATGAGACCCCTTTTGTCTCTCTCTGTCATCTGAGCCCTGTGATA 600  
Qy 185 ----- 185  
Db 601 CGCTCCTCTGCAGAAATATGGACTTGAAGGAGCAGCCAGGTAACTATTTTCGGCTGGT 660  
Qy 185 ----- 185  
Db 661 CAAGAGGATTTCCCATCTGCTGCTGTGAAGTCTGCTCTCTCTCTCTCTCTCTCTCTCT 720  
Qy 185 ----- 185  
Db 721 CTCTCAGCCCTCTCTTTCAAGAAACATGAATACCTTGGTAATTTGTCAACAGTATTACCC 780  
Qy 185 ----- 185  
Db 781 ACTGAAGAACACTTCAAGAAATGTTCAGTGAAGCTTCTAAAGAGGTCTCAGAGAGGCA 840  
Qy 185 ----- 185  
Db 841 AAAAATCTACTATAGATAGAGATTTAAACAGAGTTTTCAGAAATTAAGAAATCTCAGAAATG 900  
Qy 185 ----- 185  
Db 901 GGATCATCTGTTCACTCTCTCAAGAGCAGAAATCTGCCGTAAATAGTAGCAAAATCCTAGG 960  
Qy 185 ----- 185  
Db 961 GAAGAAATANTCGTGAAGAAATGAAGATGAAGAGAGAGTGTAGTAAATACATCCTT 1020  
Qy 185 ----- 185  
Db 1021 CATAATCAACAGAGTTACCTACAGCTCTTACTAAATTTGTTAAAGAGGATGAAGTTGTG 1080  
Qy 185 ----- 185  
Db 1081 TCTTCAGAAAGCAAAAGACAGTTTTTAATGAAGAGAGTGTGAGTGGAGCTCCCTATG 1140  
Qy 185 ----- 185  
Db 1141 AGGGAGGATATGCAGACTTCAACCAATTTGAGCGAGTATGGGAGTGAAGATAGTAAG 1200  
Qy 185 ----- 185  
Db 1201 GAAGATAGTATGTGTGCTGCTGGAGGTAAATCGAGAGCACTTGGAAAGTAAAGTG 1260  
Qy 185 ----- 185  
Db 1261 GATAAAAAATGTTTGCAGATAGCCTTGAGCAAACTAATCACGAAAAAGATAGTGAGAT 1320  
Qy 185 ----- 185  
Db 1321 AGTAATGATGATACTTCTTCCCAGTACGCCAGAGGTATAAAGGATCGTCCAGGAGCA 1380  
Qy 185 ----- 185  
Db 1381 TATATCACATGTCTCTCTTTAACCCAGCAGCACTGAGAGCATTTGCAACAAACATTTT 1440  
Qy 185 ----- 185  
Db 1441 CCTTTGTAGGAGATCCTACTTCAAGAAATAGACCCGATGAAAAAATAAGAGAAAG 1500  
Qy 185 ----- 185  
Db 1501 AAGGCCCCAATAGTAACAGAGAGAAATACTAGCACCAAAACATCAAAACCTTTTCTTGTA 1560  
Qy 185 ----- 185  
Db 1561 GCAGCACAGGATTTCTGAGACAGATTATGTCAACACAGATAATTTTAAACAAAGGTGACTGAG 1620  
Qy 185 ----- 185  
Db 1621 GAAGTCGTGGCAAAACATGCTGAGAGCCTGACTCCAGATTTAGTACAGGAAGCATGTGAA 1680

Qy 185 ----- 185  
Db 1681 AGTGAATTTGAATGAAGTTACTGGTACAAAGATTGCTTTATGAACAAAAATGGACTTGGTT 1740  
Qy 185 ----- 185  
Db 1741 CAAACATCAGAGTTATGCAAGAGTCACTCTATCTGACGACAGCTTTGCCCATCATTT 1800  
Qy 185 ----- 185  
Db 1801 GAAGAGTCAGAGCTACTCTTCCACAGTTTTCGCTGACATTGTTATGGAAGCACCATTG 1860  
Qy 185 ----- 185  
Db 1861 AATTCTCAGTTCTTAGTGTGCTTCCGTGATACAGCCAGCTCATCACCATTAGAA 1920  
Qy 185 ----- 185  
Db 1921 GCTTCTTCAGTTAATTATGAAGCATAAAAATCATGAGCCTGAAAAACCCCCACCATATGAA 1980  
Qy 185 ----- 185  
Db 1981 GAGGCCATGAGTGTATCACTAAAAAAGTATCAGGAATAAAGGAAGAAATTAAGAGCCT 2040  
Qy 185 ----- 185  
Db 2041 GAAAAATTAATGCAGCTCTTCAAGAAACAGAGCTCCTTATATCTATTTGCAATGTAT 2100  
Qy 185 ----- 185  
Db 2101 TTAATTAAGAAAAAAGCTTTCTGCTGAACAGCTCCGGATTTCTCTGATTATTTCAGAA 2160  
Qy 185 ----- 185  
Db 2161 ATGGCAAAAGTTGAACAGCAGCTGCTGATCATTTCTGAGCTAGTTGAAGATTCTCTACCT 2220  
Qy 185 ----- 185  
Db 2221 GATTTCTGAACAGCTTGACTTTATTTAGTGATGATTCAATACCTGACGTTCCACAAAAACAA 2280  
Qy 185 ----- 185  
Db 2281 GATGAACCTGTGATGCTGTGGAAGAAAGTCTCACTGAGATTTCAITTTGAGTCAATGATA 2340  
Qy 185 ----- 185  
Db 2341 GAATATGAATAAAGAAAAAACTCAGTGTCTTTGCCACCTGAGGGAGAAAGCCATATTTG 2400  
Qy 185 ----- 185  
Db 2401 GAATCTTTTAAGCTCAGTTTAGATAACACAAAAAGATACCCTGTTACCTGATGAAGTTTCA 2460  
Qy 185 ----- 185  
Db 2461 ACATTGAGCAAAAAGGAGAAAAATTCCTTTGAGATGGAGAGCTCAGTACTGCAGTTTAT 2520  
Qy 185 ----- 185  
Db 2521 TCAATGATGACTTATTTATTTCTAAGGAAGCAAGATAAGAGAAACTGAACGTTTTC 2580  
Qy 185 ----- 185  
Db 2581 GATTCATCTCCAATTGAATATATAGATGAGTTCCCTACATTGATCAGTTCTAAAAACGTAT 2640  
Qy 185 ----- 185  
Db 2641 TCATTTTCTAAATTAGCCAGGGAATATACTGACCTAGAGTATCCACAAAAAGTGAAT 2700  
Qy 185 ----- 185  
Db 2701 GCTAATGCCCCGGATGAGCTGGGTCAITGCTTTGACAGAAATTTGCCCATGACCTTTCT 2760

```

Qy 185 ----- 185
Db 2761 TTGAAGAACATACAAACCCAAAGTTGAAGAGAAAATCAGTTTCTCAGATGACTTTCTTCTAAA 2820
Qy 185 ----- 185
Db 2821 AATGGGTCTGTACATCAAAAGGTGCTCTTATTGGCTCCAGATGTTTCTGTTGGCCACT 2880
Qy 185 ----- 185
Db 2881 CAAGCAGAGATAGAGAGCATAGTTAAACCCAAAGTTCTTTGTGAAGAGAGCTGAGAAAAA 2940
Qy 185 ----- 185
Db 2941 CTTCTTCGATACAGAAAAAGAGGAGACATCACCATCTGCTATATTTTCAGCAGAGCTG 3000
Qy 186 -----ValValAspLeuLeuTyrTrpArgAspIleIysIysThrGlyValVal 201
Db 3001 AGTAAAACTTCAGTTGTTGACCTCTGTACTGGAGAGACATTAAAGAAAGACTGGAGTGGTG 3060
Qy 202 PheGlyAlaSerLeuPheLeuLeuSerLeuThrValPheSerIleValSerValThr 221
Db 3061 TTTGTGTCAGGCTATTCTGCTGCTTTTCATTGACAGATTTCAGCATTTGAGCCGTAA 3120
Qy 222 AlaTyrIleAlaLeuAlaLeuSerValThrIleSerProArgIleTyrLysGlyVal 241
Db 3121 GCCTACATTGCTTGGCCCTGCTCTCTGTGACCATCAGCTTTAGGATATACAGGGGTG 3180
Qy 242 IleGlnAlaIleGlnIysSerAspGluGlyHisProPheArgAlaTyrLeuGluSerGlu 261
Db 3181 ATCCAAAGCTATCCAGAAATCAGATGAAGGCCACCCATTTCAGGGCATATCTGGAATCTGAA 3240
Qy 262 ValAlaIleSerGluLeuValGlnIysTyrSerAsnSerAlaLeuGlyHisValAsn 281
Db 3241 GTTGCTATATCTCAGGAGTTGGTTCAGAAAGTACAGTAAATCTGCTCTTGGTCATGTGAAC 3300
Qy 282 CysThrIleLysGluLeuArgArgLeuPheLeuValAspAspLeuValAspSerLeuLys 301
Db 3301 TGCACGATAAAGAACTCAGGGCCCTCTTCTTAGTTGATGATTTAGTTGATTTCTCTGAAG 3360
Qy 302 PheAlaValLeuMetTrpValPheThrTyrValGlyAlaLeuPheAsnGlyLeuThrLeu 321
Db 3361 TTTGCAGTGTGATGGGTATTACCTATGTTGGTGCCTTGTAAATGGTCTGACACTA 3420
Qy 322 LeuIleLeuAlaIleSerLeuPheSerValProValIleTyrCluArgHisGlnAla 341
Db 3421 CTGATTTTGGCTCTCATTTTCATCTTCAGTGTCCCTGTTATTATTATGAACGGCATCAGGCG 3480
Qy 342 GlnIleAspHisTyrLeuGlyLeuAlaAsnLysAsnValLysAspAlaMetAlaLysIle 361
Db 3481 CAGATAGATCATTAICTAGGACTTGGCAATAGAAATGTTAAAGATGCTATGGCTAAATC 3540
Qy 362 GlnAlaLysIleProGlyLeuLysArgLysAlaGlu 373
Db 3541 CAAGCAAAAAATCCCTGGATTGAAGCGCAAAAGCTGAA 3576

```

Search completed: August 3, 2005, 23:25:47  
Job time : 606 secs

GenCore version 5.1.6  
 Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 3, 2005, 17:13:53 ; Search time 3100 Seconds  
 (without alignments)

4579.991 Million cell updates/sec

Title: US-09-544-776-2

Perfect score: 1910

Sequence: 1 MEDDQSLVSSSDSPRPQ.....VKDAMAKIOAKIPGLRRKAE 373

Scoring table:

BLOSUM62  
 Xgapop 10.0 , Xgapext 0.5  
 Ygapop 10.0 , Ygapext 0.5  
 Fgapop 6.0 , Fgapext 7.0  
 Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp  
 -Q=/cgp2\_1/USPTO.spool\_p/US09544776/runat\_03082005\_181345\_28214/app.query.fasta\_1.519  
 -DB=EST -QPMI=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
 -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
 -USER=US09544776@cgn 1.1 5180 @runat\_03082005\_181345\_28214 -NCPU=6 -ICPU=3  
 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
 -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:  
 1: gb\_est1:  
 2: gb\_est2:  
 3: gb\_hic:  
 4: gb\_est3:  
 5: gb\_est4:  
 6: gb\_est5:  
 7: gb\_est6:  
 8: gb\_g881:  
 9: gb\_g882:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1532.5	80.2	3533	AK034902	Mus muscu
2	1281	67.1	1097	EX439214	EX439214
3	1206	63.1	911	AL549191	AL549191
4	1138.5	59.6	781	BI079496	BI079496
5	1050	55.0	677	CN791158	CN791158
6	999.5	52.3	1002	EX462171	EX462171
7	998	52.3	924	BU845601	BU845601
8	989.5	51.8	708	BI157842	BI157842
9	981.5	51.4	810	BI080232	BI080232

10	948.5	49.7	592	7	CN482802
11	948.5	49.7	990	4	BI691132
12	945	49.5	650	6	CB215381
13	940	49.2	585	2	CO259245
14	937.5	49.1	815	2	BF099705
15	936.5	49.0	958	4	BM001698
16	932.5	48.8	679	4	BI149602
17	928	48.6	598	7	CF118424
18	925	48.4	712	7	CK971318
19	923	48.3	667	7	CN429712
20	917	48.0	600	9	AY404970
21	917	48.0	672	7	CK977984
22	917	48.0	682	6	CB162885
23	917	48.0	743	6	CD102817
24	917	48.0	758	4	BG697436
25	917	48.0	788	1	AL533461
26	917	48.0	843	4	BG570231
27	917	48.0	849	7	CR765672
28	917	48.0	875	1	AL573494
29	917	48.0	1540	3	CR611869
30	917	48.0	1785	3	AF077050
31	914	47.9	730	1	AU297347
32	911	47.7	634	6	CB067821
33	911	47.7	670	7	CV030029
34	911	47.7	960	7	CN646472
35	911	47.7	983	7	CN803408
36	911	47.7	1031	7	CN847521
37	905	47.4	805	7	CO735185
38	904	47.3	757	4	BG715173
39	903	47.3	1042	7	CN805577
40	902	47.2	871	6	CD110203
41	902	47.2	882	5	BX426505
42	901	47.2	600	9	AY404972
43	898	47.0	1013	7	CO048918
44	898	47.0	1018	7	CN801888
45	895	46.9	779	7	CO401465

#### ALIGNMENTS

#### RESULT 1

#### AK034902

#### LOCUS

#### DEFINITION

#### ACCESSION

#### VERSION

#### KEYWORDS

#### SOURCE

#### ORGANISM

#### REFERENCE

#### AUTHORS

#### TITLE

#### JOURNAL

#### MEDLINE

#### PUBMED

#### REFERENCE

#### AUTHORS

AK034902 3533 bp mRNA linear HTC 03-APR-2004  
 Mus musculus 12 days embryo embryonic body between diaphragm region  
 and neck cDNA, RIKEN full-length enriched library, clone:9430059L06  
 product:RETICULON 4 (NEURITE OUTGROWTH INHIBITOR) (NOGO PROTEIN)  
 (FOOCEN) (GLT4 VESICLE 20 KDA PROTEIN) homolog [Rattus  
 norvegicus], full insert sequence.

AK034902

AK034902.1 GI:26084268

HTC; CAP trapper.

Mus musculus (house mouse)

Mus musculus

1

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,

Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

High-efficiency full-length cDNA cloning

Meth. Enzymol. 303, 19-44 (1999)

99279253

10349636

2

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,

Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new genes

Genome Res. 10 (10), 1617-1630 (2000)

20499374

11042159

3

Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,

Konno, H., Akiyama, J., Nishi, K., Kitsuunai, T., Tashiro, H., Itoh, M.,

Sumi, N., Ighii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,

Sumi, N., Ighii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,

Sumi, N., Ighii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,

Sumi, N., Ighii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,

Sumi, N., Ighii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,

Sumi, N., Ighii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,





QY 340 GlnAlaGlnIleAspHisTyrLeuGlyLeuAlaAsnLysAsnVallyAspAlaMetAla 359  
 Db 1202 CAGCGCAGATAGATCATATTCTAGGACTTGCAACCAAGAGCGTTAAGGATGCATGGCC 1261  
 QY 360 LysIleGlnAlaValIleProGlyLeuLysArgLysAlaGlu 373  
 Db 1262 AAAATCCAGCAAAATATCCCTGGATTGAAGCGCAAGACGAA 1303

## RESULT 2

LOCUS BX439214 1097 bp mRNA linear EST 04-MAY-2004  
 DEFINITION BX439214 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA clone CS0DE008Y304  
 5-PRIME, mRNA sequence.

ACCESSION BX439214

VERSION BX439214.2 GI:47020895

KEYWORDS EST.

SOURCE Homo sapiens (human)

## ORGANISM

Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1097)

AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.

TITLE Full-length cDNA libraries and normalization

JOURNAL Unpublished (2001)

COMMENT On May 15, 2003 this sequence version replaced gi:30787776.

Contact: Genoscope

Genoscope - Centre National de Sequencage

2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized. Library was constructed by Life Technologies, a division of Invitrogen.

This sequence belongs to sequence cluster 1423.r

For more information about this cluster, see

http://www.genoscope.cns.fr/cdna?s=CS0DE008Y304p1423.r.

## FEATURES

## source

Location/Qualifiers  
 1..1097  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0DE008Y304"  
 /tissue\_type="PLACENTA"  
 /clone\_lib="Homo sapiens PLACENTA"  
 /notes="Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

## ORIGIN

## Alignment Scores:

Pred. No.: 3,86e-73 Length: 1097  
 Score: 1281.00 Matches: 256  
 Percent Similarity: 87.97% Conservative: 0  
 Best Local Similarity: 87.97% Mismatches: 35  
 Query Match: 67.07% Indels: 0  
 DB: 5 Gaps: 0

US-09-544-776-2 (1-373) x BX439214 (1-1097)

QY 1 MetGluAspLeuAspGlnSerProLeuValSerSerAspSerProArgProGln 20

Db 212 ATGAGAGCTTGACACGATCTCTCTGGTCTCTCGGACAGCCACCCGCGCGAG 271

QY 21 ProAlaPheLysTyrGlnPheValArgGluProGluAspGluGluGluGlu 40

Db 272 CCGCGCTTCAAGTACCAGTTCGTGAGGAGCCCGGAGGACGAGGAGGAGGAG 331

QY 41 GluGluGluAspGluAspGluGluGluGluGluGluGluGluGluGlu 60

Db 332 GAAGAGGAGGACGAGGACGAGGACCTGGAGGAGCTGGAGTGTCTGGAGGAGCCCGCC 391

QY 61 AlaGlyLeuSerAlaAlaProValProThrAlaProAlaAlaGlyAlaProLeuMetAsp 80  
 Db 392 GCCGGGCTGTCCGCGGCCCCAGTGGCCACCGCCCTGCCGCGCGCGCCCTGATGGAC 451  
 QY 81 PheGlyAsnAspPheValProProAlaProArgGlyPheLeuProAlaAlaProProVal 100  
 Db 452 TTCGGAATGACTTCTGTGCGCGCGGCCCGCGGGACCCCTGCGCGCGCTCCCGCGTC 511  
 QY 101 AlaProGluArgGlnProSerTrpAspProSerProValSerSerThrValProAlaPro 120  
 Db 512 GCCCGGAGCGGAGCGCGCTTCTGGACCCGAGCCCGGTCTGTCGACCGTGGCCCGGCCA 571  
 QY 121 SerPheLeuSerAlaAlaValSerProSerLysLeuProGluAspAspGluProPro 140  
 Db 572 TCCCGCTGTCTGCTCGCGCAGTCTCGCCCTCCAGCTCCCTGAGGACGAGCGCTCCG 631  
 QY 141 AlaArgProProProProProAlaSerValSerProGlnAlaGluProValTrpThr 160  
 Db 632 GCCCGCYTCCCGCTYNDWHMMAAAAAAATRAAAMCCCGAGCAGAGCCCGTGTGSSCC 691  
 QY 161 ProProAlaProAlaProAlaProProSerThrProAlaAlaProLysArgGly 180  
 Db 692 CGCCACCCCGCGCDDCCCGSGSVCCCTCCMCCCGCGCGGCCCAAGCGCAGGSGS 751  
 QY 181 SerSerGlySerValValValAspLeuLeuTyrTrpArgAspIleLysLysThrGlyVal 200  
 Db 752 TCCTCGGSGTCAAGGKGTGKACCTCTGTACTGCGAGACATTAAGAAGACTGGAGTG 811  
 QY 201 ValPheGlyAlaSerLeuPheLeuLeuLeuSerLeuThrValPheSerIleValSerVal 220  
 Db 812 GTGTTTGGTGCCAGCCTATTCTCTGCTGCTTTTCATTGACAGTATTCAGCATTTGTGAGCGTA 871  
 QY 221 ThrAlaTyrIleAlaLeuAlaLeuLeuSerValThrIleSerProArgIleTyrLysGly 240  
 Db 872 ACAGCTCATTTGCTTGGCCCTGCTCTGTGACCATCAGCTTTAGGATATACAAAGGT 931  
 QY 241 ValIleGlnAlaIleGlnLysSerAspGluGlyHisProPheArgAlaTyrLeuGluSer 260  
 Db 932 GTGATCCAAGTATCCAGAAATCAGATGAAGGCCACCCATTCAGGCGCATATCTGGAATCT 991  
 QY 261 GluValAlaIleSerGluGluLeuValGlnLysTyrSerAsnSerAlaLeuGlyHisVal 280  
 Db 992 GAAGTGTCTATATCTGARGAGTTGGTTTCAGAAATACARTTAWTTTCTCTCTGTCATKTT 1051  
 QY 281 AsnCyThrIleLysGluLeuArgLeuPhe 291  
 Db 1052 GAATKSMCGAAAAAARGAAHYARGSGCCYTTTTT 1084

## RESULT 3

## ALS49191

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## CONTACT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

was normalized. Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 1423.r

For more information about this cluster, see  
http://www.genoscope.cns.fr/cdna?c=CSODI051CA06QP1&c=1423.r.

## FEATURES

## source

```

1. 911
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODI051YB11"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

```

## ORIGIN

Alignment Scores:  
Pred. No.: 2,2e-68 Length: 911  
Score: 1206.00 Matches: 243  
Percent Similarity: 93.85% Conservative: 1  
Best Local Similarity: 93.46% Mismatches: 15  
Query Match: 63.14% Indels: 1  
DB: 1 Gaps: 0

US-09-544-776-2 (1-373) x AL549191 (1-911)

```

QY 1 MetGluAspLeuAspGlnSerProLeuValSerSerAspSerProProArgProGln 20
DB 134 ATGGAGACTGGACCAAGTCTCTCTGCTCTGCTCTGGACAGCCACCCCGCGCGAG 193
QY 21 ProAlaPheLysTyGlnPheValArgGluProGluAspGluGluGluGluGlu 40
DB 194 CCGCGGCTTCAAGTACCAGTTCGTGAGGAGCCGAGGACGAGGAGGAGGAGGAG 253
QY 41 GluGluGluAspGluAspGluAspLeuGluGluLeuGluValLeuGluArgLysProAla 60
DB 254 GAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 313
QY 61 AlaGlyLeuSerAlaAlaProValProThrAlaProAlaProAlaGlyAlaProLeuMetAsp 80
DB 314 GCGGGCTCTCGCGGCCCCAGTGGCCACCGCCCCCTGCGCGCGCGCGCTCTGATGAC 373
QY 81 PheGlyAsnAspPheValProProAlaProAlaProArgGlyPheLeuProAlaAlaProProVal 100
DB 374 TTCCGAAATGACTTCGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTCCCGCCGTC 433
QY 101 AlaProGluArgGlnProSerTrpAspProSerProValSerSerThrValProAlaPro 120
DB 434 GCGCGGAGCGGAGCGGCTTGTGGAGCCGAGCCGCGGTCTGTMGACCGTGGCCGCCA 493
QY 121 SerPheLeuSerAlaAlaAlaValSerProSerLysLeuProGluAspAspGluProPro 140
DB 494 TCCCGCTGTCTGCTGCGCGAGTCTCGCCCTCCAAAGTCCCTGAGGACGACGAGCTNCG 553
QY 141 AlaArgProProProProProAlaSerValSerProGlnAlaGluProValTrpThr 160
DB 554 GNCGGGCHTMMHCCTAADYACMMWGGCAGCGTGAACCCCGCGAGCGCGCTGTGGMCC 613
QY 161 ProProAlaProAlaProAlaProAlaProSerThrProAlaAlaProLysArgGln 180
DB 614 CCGSAGCCCGGCTCSCCGCGGCCCCCTCAACCCCGCGCGCGCCCMWCGCAGGGG 673
QY 180 ySerSerGlySerValValValAspLeuLeuTyTrpArgAspIleLysLysThrGlyVa 200
DB 674 CTCCTCGGCTCAGTGTGTTGTTGACCTCTGTACTGGAGAGACATTAAGAGACTGGAGT 733
QY 200 lValPheGlyAlaSerLeuPheLeuLeuSerLeuThrValPheSerIleValSerVa 220
DB 734 GGTGTTTGTGTCCAGCCTATTCTGCTGCTTTTACAGCAGTATTACAGCATTTGTGAGCGT 793

```

```

QY 220 lThrAlaTyrlleAlaLeuAlaLeuLeuSerValThrIleSerProArgIleTyrlLysG 240
DB 794 AACAGCCATACATTGCTTGGCCCTCTCTCTGTGACCATCAGCTTTAGGATATACAAGG 853
QY 240 yValIleGlnAlaIleGlnLysSerAspGluGlyHisProPheArgAlaTyrlLeuGlu 259
DB 854 TGTGATCCAGCTATCCAGAAATCAGATGAAGCCCACTTCAGGGCATATCTGGAA 911

```

## RESULT 4

## BI079496

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## FEATURES

## source

## Location/Qualifiers

## 1. 781

## /organism="Mus musculus"

## /mol\_type="mRNA"

## /strain="FVB/N-3"

## /db\_xref="taxon:10090"

## /clone="IMAGE:5008248"

## /tissue\_type="tumor, biopsy sample"

## /dev\_stage="5 months"

## /lab\_host="DH10B"

## /clone\_lib="NCI\_CGAP\_Mam2"

## /note="organ: mammary; Vector: pCMV-SPORT6; Site\_1: SalI; Site\_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies. Investigator providing samples: Gilbert Smith, NIH"

## ORIGIN

## Alignment Scores:

## Pred. No.:

## Score:

## Percent Similarity:

## Best Local Similarity:

## Query Match:

## DB:

## 4.27e-64

## 1138.50

## 89.36%

## 87.55%

## 59.61%

## 4

## Length:

## Matches:

## Conservative:

## Mismatches:

## Indels:

## Gaps:

## 781

## 239

## 5

## 14

## 15

## 4

## US-09-544-776-2 (1-373) x BI079496 (1-781)

## QY

## 101

## AlaProGluArgGlnProSerTrpAspProSerProValSerSerThrValProAlaPro

## DB

## 3

## GCCCTGAGAGGACGCCGTCTCTGGAAACGACGCCCGCGCGCTCC-----GGCCCA

## QY

## 121

## SerPheLeuSerAlaAlaAlaValSerProSerLysLeuProGluAspAspGluProPro

## DB

## 54

## TCCCTGCCCGCTGCTGCCGAGTCTGCTCCCAAGCTCCCGAGGACGACGAGCTCCA

## QY

## 141

## AlaArgProProProProProAlaSerValSerProGlnAlaGluProValTrpThr

## DB

## 114

## GCGCGG-----CCTCCGCGCGCGAGCGCGCGCGCTAGCGAG-----

161 ProProAlaProAlaAlaProProSerThrProAlaAlaProProLysArgArgGly 180  
 156 -----CCGCGCGCGCCCTCCACGCGCGCGCGCCCAAGCGCAGGGC 200  
 181 SerSerGlySerValValValAspLeuLeuTyrTrpArgAspLeuLysThrGlyVal 200  
 201 ----TCGGGCTCAGTGGTGTGTACCTCTCTACTGGAGAGACATTAGAAGACTGGAGTG 257  
 201 ValPheGlyAlaSerLeuPheLeuLeuSerLeuThrValPheSerIleValSerVal 220  
 258 GTGTTTGTCGACGCTTATTCCTGCTGCTGCTCTGACAGTTCACATGTTCAGTGTA 317  
 221 ThrAlaTyrIleAlaLeuAlaLeuLeuSerValThrIleSerProArgIleTyrLysGly 240  
 318 ACGGCTTACATTGCTTTGGCCCTGCTCTCTGTGACTATCAGCTTTAGGATATATAAGGCT 377  
 241 ValIleGlnAlaIleGlnLysSerAspGluGlyHisProPheArgAlaTyrLeuGluSer 260  
 378 GTGATCCAAAGCTATCCAGAAATCAGATGAAGGCCACCCATTCAAGGCATATTTGGAATCT 437  
 261 GluValAlaIleSerGluGluLeuValGlnLysTyrSerAsnSerAlaLeuGlyHisVal 280  
 438 GAAGTTGCCATATCAGAGGAATGGTTCAGAAATATAGTAATCTGCTCTTGGTCATGTG 497  
 281 AsnCysThrIleLysGluLeuArgArgLeuPheLeuValAspLeuValAspSerLeu 300  
 498 AACAGCACATAAAGAAATGAGCGCTCTCTCTAGTTCATGATTAGTTGATTCCTCG 557  
 301 LysPheAlaValLeuMetTrpValPheThrTyrValGlyAlaLeuPheAsnGlyLeuThr 320  
 558 AAGTTTGCAGTGTGTGATGGTATTTACTTACGTTGGTGCCTTGTTCATGTTGACA 617  
 321 LeuLeuIleLeuAlaLeuIleSerLeuPheSerValProValIleTyrGluArgHisGln 340  
 618 CTACTGATTTAGTCTGATCTCAGTCTCAGTCTTCTGATTTCTCTTATATATGAACGGATCAG 677  
 341 AlaGlnIleAspHisTyrLeuGlyLeuAlaAsnLysAsnValLysAspAlaMetAlaLys 360  
 678 GCCAGATCGATCATTTATCTAGGCTTGCAACAGAGCGTTAAGGATGCCATGCGCAA 737  
 361 IleGlnAlaLysIleProGlyLeuLysArgGlyAlaGlu 373  
 738 ATCCAGCAAAATCCCTGGATTGAAGCGCAAGCAGAT 776

RESULT 5  
 CN791158  
 LOCUS CN791158 677 bp mRNA linear EST 26-MAY-2004  
 DEFINITION 4125836 BARC 8BOV Bos taurus cdna clone 8BOV\_41013 5', mRNA  
 ACCESSION CN791158  
 VERSION CN791158.1 GI:47687138  
 KEYWORDS EST.  
 SOURCE Bos taurus (cow)  
 ORGANISM Bos taurus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 Bovinae; Bos.  
 1 (bases 1 to 677)  
 Baumann,R.G.; Baldwin,R.L.; Sonstegard,T.S.; Van Tassel,C.P. and  
 Matukumalli,L.K.  
 TITLE Construction and Analysis of a cDNA Library Generated From  
 Intestinal Muscle and Epithelial Tissues of Holstein Cattle  
 JOURNAL Unpublished (2004)  
 COMMENT Contact: Richard G. Baumann  
 Bovine Functional Genomics Lab  
 ANRI  
 BLDG 162; BARC-EAST, Beltsville, MD 20705, USA  
 Tel: 3015048604  
 Fax: 3015048744  
 Email: rbaumann@anri.barc.usda.gov  
 Single pass sequencing. Bases called and trimmed with phred  
 0.00925 using options -trim\_alt'' -trim\_fasta. Vector identified

by cross match using options -minmatch 12 -minscore 18  
 Plate: 41 row: 0 column: 13  
 Seq primer: CCTATTAGTGACACTATAGAAC  
 High quality sequence stop: 677.

# FEATURES

source  
 1..677  
 /organism="Bos taurus"  
 /mol\_type="mRNA"  
 /strain="Holstein"  
 /db\_xref="taxon:9913"  
 /clone="8BOV\_41013"  
 /sex="female"  
 /tissue\_type="Epithelial, Muscle"  
 /dev\_stage="Lactating, Neonatal"  
 /lab\_host="DH10B Tona"  
 /clone\_lib="BARC 8BOV"  
 /note="Organ: Intestine; Vector: pCMVSPORT6.1; Site\_1:  
 NotI; Site\_2: EcoRI; Normalized cow cDNA intestinal  
 library in pCMVSPORT6.1, constructed from equimolar mRNA  
 pools derived from 5 sources, 4 lactating intestinal, 1  
 neonatal intestinal 4/5 Lactating, Proximal Duodenum,  
 Jejunum, Distal Ileum, Colon, 1/5 Neonatal, Proximal  
 Duodenum, Jejunum, Distal Ileum"

# ORIGIN

Alignment Scores:  
 Pred. No.: 1..92e-58 Length: 677  
 Score: 1050.00 Matches: 213  
 Percent Similarity: 96.86% Conservatives: 3  
 Best Local Similarity: 95.52% Mismatches: 7  
 Query Match: 54.97% Indels: 1  
 DB: 7 Gaps: 0

US-09-544-776-2 (1-373) x CN791158 (1-677)

QY 137 AspGluProProAlaArgProProProProProProAlaSerValSerProGlnAlaGlu 156  
 Db 10 GAGGAGCCTCCGCGCGCGCTCCCTCTCTCA-CCGCGCGACGTGACCCCTGAGCGCGAG 68  
 QY 157 ProValTyrProProAlaProAlaProAlaProProSerThrProAlaAlaPro 176  
 Db 69 CCCTCGTGACCTCGGTGTGTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGACCC 128  
 QY 177 LysArgArgGlySerSerGlySerValValValAspLeuLeuTyrTrpArgAspIleLys 196  
 Db 129 AAAGCAGCGGCTCTCTCGGCTCAGTGGTGTGACCTCTCTTACTGGAGACATTAAAG 188  
 QY 197 LysThrGlyValValPheGlyAlaSerLeuPheLeuLeuLeuSerLeuThrValPheSer 216  
 Db 189 AAGACTGGAGTGGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 248  
 QY 217 IleValSerValThrAlaTyrIleAlaLeuAlaLeuLeuSerValThrIleSerProArg 236  
 Db 249 ATTGTGAGTGTAAAGCGGCTTACATTTGCGGCTCTCTCTGTGACTATCAGCTTTAGG 308  
 QY 237 IleTyrLysGlyValIleGlnAlaIleGlnLysSerAspGluGlyHisProPheArgAla 256  
 Db 309 ATATATAAGGCTGTGATCCAGGCTATCCAGAAATCTGATGAAGCGCCACCCATTGAGGCA 368  
 QY 257 TyrLeuGluSerGluValAlaIleSerGluGluLeuValGlnLysTyrSerAsnSerAla 276  
 Db 369 TATTTGGAATCTGAAGTTCATATCTGAGGAGTGGTTCAGAGGTACAGCAATTCGTCT 428  
 QY 277 LeuGlyHisValAsnCysThrIleLysGluLeuArgGlyPheLeuValAspAspLeu 296  
 Db 429 CTTGGTGTGTTAACTGCACAAATAAAGAACTCAGACGCGCTCTCTTAGTGTGATGATTTA 488  
 QY 297 ValAspSerLeuLysPheAlaValLeuMetTrpValPheThrTyrValGlyAlaLeuPhe 316  
 Db 489 GTTGATTTCTGAAGTTTCAGTGTGGTATTTACCTATCTTGGTGGCTTGTTC 548  
 QY 317 AsnGlyLeuThrLeuLeuLeuAlaLeuIleSerLeuPheSerValProValIleTyr 336

Db 549 AATGGTCTGACACTACTAATTTTGGCTCTGATTTTCACCTCTTCAGTGTCTCTGTTATTAT 608  
 QY 337 GluArgHisGlnAlaGlnIleAspHisTyrLeuGlyLeuAlaAsnLysenValIysAsp 356  
 Db 609 GAACGGCATCAGCGCAATAGATCATTTATCTGGGACTTGGCAATTAAGAATGTTAAAGT 668  
 QY 357 AlaMetAla 359  
 Db 669 GCTATGGCT 677

RESULT 6  
 BX462171 1002 bp mRNA linear EST 06-MAY-2004  
 LOCUS BX462171 Homo sapiens B CELLS (RAMOS CELL LINE) Homo sapiens cDNA  
 DEFINITION Clone CS0DG002YP16 5-PRIME, mRNA sequence.  
 ACCESSION BX462171  
 VERSION BX462171.2 GI:47071945  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 1002)  
 Li.W.B., Gruber.C., Jessee.J. and Polayes.D.  
 Full-length cDNA libraries and normalization  
 Unpublished (2001)  
 On May 22, 2003 this sequence version replaced gi:31035379.  
 Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 2 rue Gaston Cremlieux, CP 5706 - 91057 EVRY cedex - FRANCE  
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
 end enriched, double-strand cDNA was digested with Not I and cloned  
 into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library  
 was not normalized. Library was constructed by Life Technologies, a  
 division of Invitrogen.

This sequence belongs to sequence cluster 1423.r  
 For more information about this cluster, see  
 http://www.genoscope.cns.fr/cdna?c=CS0DG002DH08QPlsc=1423.r.  
 FEATURES  
 source  
 1. .1002  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0DG002YP16"  
 /tissue\_type="B CELLS (RAMOS CELL LINE)"  
 /cell\_line="RAMOS CELL LINE"  
 /clone\_lib="Homo sapiens B CELLS (RAMOS CELL LINE)"  
 /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed  
 with a NotI-oligo(dT) primer. Five prime end enriched,  
 double-strand cDNA was digested with Not I and cloned into  
 the Not I and EcoRV sites of the pCMVSPORT 6 vector.  
 Library was not normalized."

## ORIGIN

Alignment Scores:  
 Pred. No.: 5,75e-55 Length: 1002  
 Score: 999.50 Matches: 223  
 Percent Similarity: 76.00% Conservative: 5  
 Best Local Similarity: 74.33% Mismatches: 46  
 Query Match: 52.33% Indels: 28  
 DB: 5 Gaps: 4

US-09-544-776-2 (1-373) x BX462171 (1-1002)

QY 1 MetGluAspLeuAspGlnSerProLeuValSerSerSerAspSerProProArgProGln 20  
 Db 145 ATGGAAGACCTGGACCAAGTCTCTCTGTCGTCTCGACAGCCACCCCGCGCGCAG 204  
 QY 21 ProAlaPheLysTyrGlnPheValArgGluPro-GluAspGluGluGluGluGluGl 40  
 Db 205 CCCCGGTTCAAGTACCAAGTTCGTGAGGGAGCCGCGAGGACGAGGAAGAAGAGGAGGA 264

QY 40 uGluGluGluAspGluAspGluAspGluGluGluValLeuGluValLeuGluArgLysProAl 60  
 Db 265 GGAAGAGGAGGACGAGGAGCAAGAGACTGGAGAGAGCTGGAGGTGCTGGAGAGAGAGCCCGC 324  
 QY 60 aAlaGlyLeuSerAlaAlaProValProThrAlaProAlaAlaGlyAlaProLeuMetAs 80  
 Db 325 CGCGGGCTGTCCGCGGCCCCAGTCCCAACACCCCTGCCCGCGCGCGCGCTGATGA 384  
 QY 80 pPheGlyAsnAspPheValProAlaProArgGlyPheLeuProAlaAlaProProVa 100  
 Db 385 CTTTCGAAATGACTTCGTCCGCGCGCGCCCGCGGACCCCTGCCCGCGCGCTCCCGCGT 444  
 QY 100 lAlaProGluArgGlnProSerTIPAspProSerProValSerSerThrValProAlaPr 120  
 Db 445 CGCCCCGAGCGGCGAGCCGTCTTGGACCCGAGCCCGGTGCTGTCGACCGTCCCGCGCGC 504  
 QY 120 oSerPheLeuSerAlaAlaValSerProSerLysLeuProGluAspAspGluProPr 140  
 Db 505 ATCCCGCTGTCTGCTCGCGCAGTCTCGCCCTCCAAAGCTCCCTGAGGACGAGCGCTCC 564  
 QY 140 oAla-----ArgProProProProProProProAlaSerVa 151  
 Db 565 GCGCCGGTTTTCTTTTCTCTTTTMAAMAWRAMMCCA----- 601  
 QY 151 lSerProGlnAlaGluProValTIPThrProProAlaProAlaProAlaProProSe 171  
 Db 602 -----AGCAAAACCGTGTGDC-CCGCCARC-CCCGCCCCCGGGSSSSSSRS 649  
 QY 171 rThrProAlaProLysArgArgGlySerGlySerValValValValValValValVal 191  
 Db 650 CCGSGGGCGCMAG-----CSMGGGGCGCCCGSYCA---GTGGTGTGACCCCGTA 697  
 QY 191 rTPRgAspIleLysLysThrGlyValValPheGlyAlaSerLeuPheLeuLeuSe 211  
 Db 698 CTGGAGACACATTAAAGACCTGGAGTGGTGTGGTCCAGCCATTCTCTCTCTCTTC 757  
 QY 211 rLeuThrValPheSerIleValSerValThrAlaTyrIleAlaLeuAlaLeuSerVa 231  
 Db 758 ATTGACAGTATTCAGCATTTGGAGGTAAACAGCTACATTGCTTGGCCCTCTCTCTGT 817  
 QY 231 lThrIleSerProArgIleTyrLysGlyValIleGlnAlalleGlnLysSerAspGluGl 251  
 Db 818 GACCATCAGCTTAGCATATACAGGGTGTGATCAAGCTATCCAGAAATCAGATGAARG 877  
 QY 251 yHisProPheArgAlaTyrLeuGluSerGluValAlaIleSerGluLeuValGlnly 271  
 Db 878 SCMCCCATTCAGGGGATATYGGGAWYTGAAATTKSTATATYTKRRGGGKTGGTCARAA 937  
 QY 271 stYrSer-AsnSerAlaLeuGlyHisValAsnCysThrIleLysGluLeuArgArg 289  
 Db 938 AWAATAAATWTTKSTCTYTTGGKCATDTTKAAWTHSACVAAWAAARAAWCAAGCG 993

RESULT 7  
 BUB45601 924 bp mRNA linear EST 16-OCT-2002  
 LOCUS AGENCOURT 10414454 NIH MGC 109 Homo sapiens cDNA clone  
 DEFINITION IMAGE:6575264 5', mRNA sequence.  
 ACCESSION BUB45601  
 VERSION BUB45601.1 GI:24030042  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 924)  
 NIH-MGC http://mgi.nci.nih.gov/.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-r@mail.nih.gov  
 Tissue Procurement: ATCC  
 cDNA Library Preparation: Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)



```
QY 166 ProAlaAProProSerThrProAlaAProLysArgGlySerSerGlySerVal 185
Db 2 CCCGCCGCCCCCTTCACAGCGCGCGCCGACAGCGAGGGC---TCGGGCTCAGTG 58
QY 186 ValValAspLeuLeuTyrTrpArgAspIleLysThrGlyValValPheGlyAlaSer 205
Db 59 GTTGTGTGACCTCTGTACTGGAGACATTAAGAAGACTGGAGTGGTGTGGTGCCAGC 118
QY 206 LeuPheLeuLeuLeuSerLeuThrValPheSerIleValSerValThrAlaTyrIleAla 225
Db 119 TTATTCCTCTGCTGCTCTGCAGGTGTTTCAGCATTTGTCAAGTGAACGCCCTACATTGCC 178
QY 226 LeuAlaLeuLeuSerValThrIleSerProArgIleTyrLysGlyValIleGlnAlaIle 245
Db 179 TTGGCCCTGCTCTCTGTGACTATCAGCTTTAGGATATATAAGGGTGTGATCAAGCTATC 238
QY 246 GlnLysSerAspGluGlyHisProPheArgAlaTyrLeuGluSerGluValAlaIleSer 265
Db 239 CAGAAATCAGATGAAGGCCACCATTCAGGGCATATTTGGAATCTGAAAGTTGCCATATCA 298
QY 266 GluGluLeuValGlnLysTyrSerIleValSerAlaLeuGlyHisValAsnCysThrIleLys 285
Db 299 GAGGAATTGGTTTCAGAAATATAGTAATTCCTGCTCTTGGTTCATGTGAACAGCAATAAAA 358
QY 286 GluLeuArgGlyPheLeuValAspLeuValAspLeuValAspLeuLysPheAlaValLeu 305
Db 359 GAATGAGCGGCTCTCTTAGTTGATGATTTAGTTGATTCCTCAAGTTTCAGATTTTGTAGCT 418
QY 306 MetTrpValPheThrTyrValGlyAlaLeuPheAsnGlyLeuThrLeuLeuIleLeuAla 325
Db 419 ATGTGGGTATTTACTTACGTTGGTGGCTTGTTCATTTGTTCAATGGTTTCACACTACTGATTTAGCT 478
QY 326 LeuIleSerLeuPheSerValProValIleTyrGluArgHisGlnAlaGlnIleAspHis 345
Db 479 CTGATCTCACTCTCAGTATTCTCTGTATATATATGAACGGCATCAGCGCGCAGATAGCAT 538
QY 346 TyrLeuGlyLeuAlaAsnLysAsnValLysAspAlaMetAlaLysIleGlnAlaLysIle 365
Db 539 TATCTAGGACTTGCMAAACAAGCGTTAAGGATGCCATGGCCCAAAATCCAGCAAAATC 598
QY 366 ProGlyLeuLysArgLysAlaGlu 373
Db 599 CCTGGATTGAAGCGCAAGACAGAA 622

RESULT 9
BI080232
LOCUS
DEFINITION BI080232 810 bp mRNA linear EST 20-JUN-2001
602876650F1 NCI_CGAP_Mam2 Mus musculus cDNA clone IMAGE:5008092 5',
mRNA sequence.
ACCESSION BI080232
VERSION BI080232.1 GI:14498562
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Place: L14W11052 row: d column: 13
High quality sequence stop: 802.
Location/Qualifiers
```

```
source
1..810
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N-3"
/db_xref="taxon:10090"
/clone="IMAGE:5008092"
/tissue_type="tumor, biopsy sample"
/dev_stage="5 months"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Mam2"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"

ORIGIN
Alignment Scores:
Pred. No.: 6,49e-54 Length: 810
Score: 981.50 Matches: 227
Percent Similarity: 86.03% Conservative: 7
Best Local Similarity: 83.46% Mismatches: 17
Query Match: 51.39% Indels: 22
DB: 4 Gaps: 4

US-09-544-776-2 (1-373) x BI080232 (1-810)
QY 107 SerTrpAspProSerProValSerSerThrValProAlaProSerPheLeuSerAlaAla 126
Db 2 TCCTGGGAACGAGCCCGCGCGGTCC-----GGCCATCCTCGCGCCCGCTGCC 52
QY 127 AlaValSerProSerLysLeuProGluAspGluProProAlaArgProProPro 146
Db 53 GCAGTCTCTCCCTCAAGCTCCCGGAGGACGAGCTCCAGCGCGG-----CCTCG 106
QY 147 ProAlaSerValSerProGlnAlaGluProValThrProProAlaProAlaPro 166
Db 107 GCGCCAGCGCGCGCGAGCCCTCAGCGAG-----CCC 139
QY 167 AlaAlaProProSerThrProAlaAlaProLysArgArgGlySerSerGlySerVal 186
Db 140 GCGCGGCCCTTCCACCGCGCGCCGCGCCGCGCGCGCGCGCGG---TCGGGCTCAGTGGT 196
QY 187 ValAspLeuLeuTyrTrpArgAspIleLysThrGlyValValPheGlyAlaSerLeu 206
Db 197 GTTGACCTCTCTACTCGGAGAGACATTAAGAAGACTGGAGTGGTGTGGTGCCAGCTTA 256
QY 207 PheLeuLeuSerSerLeuThrValPheSerIleValSerValThrAlaTyrIleAlaLeu 226
Db 257 TTCTCTGCTGCTCTCTGACAGTGTTCAGCATTTGTCAAGCGCTTACATTGCCCTTG 316
QY 227 AlaLeuLeuSerValThrIleSerProArgIleTyrLys-GlyValIleGlnAlaIleG 246
Db 317 GCGCTGCTCTCTGTGACTATCAGCTTTAGGATATATAACGGGTGTGATCCAAGCTATCCA 376
QY 246 nLysSerAspGluGlyHisProPheArgAlaTyrLeuGluSerGluValAlaIleSerGl 266
Db 377 GAAATCAGATGAAGCCACCCATTTCAGGCGCATATTTGGAATCTGAAGTTGCCATATCAGA 436
QY 266 uGluLeuValGlnLysTyrSerAsnSerAlaLeuGlyHisValAsnCysThrIleLys-G 286
Db 437 CGAATTTGGTTTCAGAAATATAGTAATTCCTGCTCTGTCATGTGAACAGCACACATAAAC 496
QY 286 luLeuArgArgLeuPheLeuValAsp-AspLeuValAspSerLeuLysPheAlaVal-Le 305
Db 497 ACTTGAAGCGCTCTCTTCTTAGTCTGACTGATTTAGTTGATTCCTGAAGTTTCAGATGTCG 556
QY 305 uMetTrpValPheThrTyrValGlyAlaLeuPheAsnGlyLeuThrLeuLeuIleLeuAl 325
Db 557 ACTGTGGGTATTTACTTACGTTGGTGC-CTGTTCAATGGTTTGACACTACTGATTTTCAGC 615
QY 325 aleuIleSerLeuPheSerValProVal-IleTyrGluArg-HisGlnAlaGlnIleAsp 344
Db 616 TCTGATCTCACTCTTTCAGTATTCTCTGTTTATATATGAACGGCCGCGCATGATAGAT 675
```





```

/dev_stage="5 months"
/lab_host="DH108"
/clone_lib="NCI_CGAP_Mam6"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigator
providing samples: Jeffrey Green, M.D., NIH"

```

# ORIGIN

Alignment Scores:  
 Pred. No.: 1,13e-51 Length: 990  
 Score: 948.50 Matches: 200  
 Percent Similarity: 97.13% Conservative: 3  
 Best Local Similarity: 95.69% Mismatches: 4  
 Query Match: 49.66% Indels: 3  
 DB: 4 Gaps: 1

US-09-544-776-2 (1-373) x B1691132 (1-990)

QY	166	ProAlaAaProSerThrProAlaAaProLysAtgArgGlySerSerGlySerVal	185
DB	7	CCCGCGCGCCCTTCACAGCGCGCGCGCGCGAG-GGC---TCGGGCTCAGTG	62
QY	186	ValValAspLeuLeuTyrTrpArgAspIleLysLysThrGlyValValPheGlyAlaSer	205
DB	63	GTGTGTGACCTCTGTACTGGAGACATTAAGAGACTGGAGTGTGTGTGGCCAGC	122
QY	206	LeuPheLeuLeuSerLeuThrValPheSerIleValSerValThrIleAla	225
DB	123	TTATTCTCTGTCTGTCTGACAGTGTTCAGCATTTGTGAGTAAAGCGCTACATTTGC	182
QY	226	LeuAlaLeuLeuSerValThrIleSerProArgIleTyrLys-GlyValIleGlnAlaTrl	245
DB	183	TTGGCCCTCTCTGTGACTATCAGCTTTAGATATATAACGGGTGTGATCCAGCTAT	242
QY	245	eGlnLysSerAspGluGlyHisProPheArgAlaTyrLeuGluSerGluValAlaIleSe	265
DB	243	CCAGAAATCAGATGAAGGCCACCCATTCAGGCGCATATTTGGAATCTGAAGTTGCCATATC	302
QY	265	rGluGluLeuValGlnLysTyrSerAsnSerAlaLeuGlyHisValAsnCysThrIleLy	285
DB	303	AGAGGAATTTGGTTTCAGAAATATAGTAATCTGTCTTGTGTATGTGAACAGCACAAATAA	362
QY	285	sGluLeuArgArgLeuPheLeuValAspLeuValAspSerLeuLysPheAlaValLe	305
DB	363	AGAAATGAGCGGCTCTCTTTAGTTGATGATTTAGTTGATTCCTCGAAGTTTGCAGTGT	422
QY	305	uMetTrpValPheThrTyrValGlyAlaLeuPheAsnGlyLeuThrLeuLeuIleLeuAl	325
DB	423	GATGTGGGTATTACTTACGTGGTGGCTTGTTCATATGTTTGACACTACTGATTTTACG	482
QY	325	alLeuIleSerLeuPheSerValProValIleTyrGluArgHisGlnAlaGlnIleAspHi	345
DB	483	TCTGATCTCACTCTTCAGTATTCTCTGTATATATGAACGGCATTCAGCGCGCATAGATCA	542
QY	345	stYrLeuGlyLeuAlaAsnLysAsnValLysAspAlaMetAlaLysIleGlnAlaLysIl	365
DB	543	TTATCTAGGACTTGCNAACAAGAGCGTTAAGATGCAATGGCCNAATCCAAAGCAAAAT	602
QY	365	eProGlyLeuLysArgLysAlaGlu	373
DB	603	CCCTGGATTGAAGCGCAAGAGCAA	627

RESULT 12  
 CB215381  
 LOCUS  
 DEFINITION NISC np05dl2.y1 NICHDS HS\_Ut1 Homo sapiens cDNA clone IMAGE:5937070  
 5', mRNA sequence.  
 ACCESSION CB215381  
 VERSION CB215381.1 GI:28263573  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 650)  
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 Unpublished (1997)  
 Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
 CDNA Library Preparation:  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium/LLNL  
 DNA Sequencing by: National Institutes of Health Intramural  
 Sequencing Center (NISC)  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
[info@image.llnl.gov](mailto:info@image.llnl.gov)  
 Plate: LLAM13163 row: G column: 23  
 Seq primer: M13RP1 reverse primer (ABI).

## FEATURES

## source

1..650  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5937070"  
 /sex="female"  
 /tissue\_type="normal endometrium, late proliferative  
 phase, cycle day 13"  
 /lab\_host="DH10B (T1-resistant)"  
 /clone\_lib="NICHDS HS\_Ut1"  
 /note="Organ: uterus; Vector: pCMV-SPORT6.1.cdb (ResGen,  
 Invitrogen Corporation); Site 1: NotI; Site 2: EcoRV;  
 Cloned unidirectionally from microquantity amounts of mRNA  
 from normal endometrial tissue (late proliferative phase,  
 cycle day 13). Average insert size 1.9 kb. Library  
 constructed by ResGen (Invitrogen Corporation)."

## ORIGIN

Alignment Scores:  
 Pred. No.: 1,14e-51 Length: 650  
 Score: 945.00 Matches: 193  
 Percent Similarity: 99.48% Conservative: 0  
 Best Local Similarity: 99.48% Mismatches: 1  
 Query Match: 49.48% Indels: 0  
 DB: 6 Gaps: 0  
 US-09-544-776-2 (1-373) x CB215381 (1-650)  
 QY 180 GlySerSerGlySerValValValAspLeuLeuTyrTrpArgAspIleLysLysThrGly 199  
 DB 1 GGCTCTCGGGCTCAGTGGTGTTCGACCTCTCTGTACTGGAGAGACATTAAAGAAGACTGGA 60  
 QY 200 ValValPheGlyAlaSerLeuPheLeuLeuSerLeuThrValPheSerIleValSer 219  
 DB 61 GTGGGTGTGGTGGCGCAGGCTATTCCTGCTGCTTTTATTGACAGATATTCAGCAATTGTGAGC 120  
 QY 220 ValThrAlaTyrIleAlaLeuAlaLeuSerValThrIleSerProArgIleTyrLys 239  
 DB 121 GTPACAGCTACATTGCCCTTGGCCCTGCTCTCTGTGACCATCAGCTTTAGGATATACAG 180  
 QY 240 GlyValIleGlnAlaIleGlnLysSerAspGluGlyHisProPheArgAlaTyrLeuGlu 259  
 DB 181 GGTGTGATCCAGCTATCCAGAAATCAGATGAAGGCCACCCATTGAGGCATATCTGGAA 240  
 QY 260 SerGluValAlaIleSerGluGluValGlnLysTyrSerAsnSerAlaLeuGlyHis 279  
 DB 241 TCTGAAGTTGCTATATCTGAGGAGTGTGGTTCAAGAGTACAGTAATTTCTCTCTTGTGTA 300  
 QY 280 ValAsnCysThrIleLysGluLeuArgArgLeuPheLeuValAspLeuValAspSer 299  
 DB 301 GTGAATCGACGATAAAGGAACCTCAGGCGCTCTTCTTAGTTGATGATTTAGTATTCT 360  
 QY 300 LeuLysPheAlaValLeuMetTrpValPheThrTyrValGlyAlaLeuPheAsnGlyLeu 319

Db 361 CTGAAGTTTCAGTGGTGGTATTTACCTATGTTGGTGGCTTTGTTTAATGTTCTG 420

Qy 320 ThrLeuLeuLeuAlaLeuLeuSerLeuPheSerValProValIleTyrGluArgHis 339

Db 421 ACACACTGATTTGGCTCTCATTTCTACTCTTCAGTGTTCCTGTTATTAAGACGGCAT 480

Qy 340 GlnAlaGlnIleAepHisTyrLeuGlyLeuAlaAenLysAenValLysAepAlaMetAla 359

Db 481 CAGGCACAGATAGATCATTTATCTAGGACTTGCAAAATGAAGATGTTAAAGATGCTATGGCT 540

Qy 360 LysIleGlnAlaLysIleProGlyLeuLysArgLysAlaGlu 373

Db 541 AAAATCCAGCAAAATCCCTGGATTGAAGCGCAAGCTGAA 582

RESULT 13

LOCUS CO259245

DEFINITION 4130644 BARC 8BOV Bos taurus cDNA clone 8BOV\_51K20 5', mRNA

ACCESSION CO259245

VERSION CO259245.1 GI:49144047

KEYWORDS EST.

SOURCE Bos taurus (cow)

ORGANISM Bos taurus

REFERENCE 1 (bases 1 to 585)

AUTHORS Baumann,R.G., Baldwin,R.L., Sonstegard,T.S., Van Tassel,C.P. and Matukumalli,L.K.

TITLE Construction and Analysis of a cDNA Library Generated From Intestinal Muscle and Epithelial Tissues of Holstein Cattle

JOURNAL Unpublished (2004)

COMMENT Contact: Richard G. Baumann  
Bovine Functional Genomics Lab  
ANRI  
BLDG 162: BARC-EAST, Beltsville, MD 20705, USA  
Tel: 3015048604  
Fax: 3015048744  
Email: rbaumann@anri.barc.ueda.gov

Single pass sequencing. Bases called and trimmed with phred 0.000925 using options -trim\_alt -, -trim\_fasta. Vector identified by cross match using options -minmatch 12 -minscore 12

Plate: 51 row: K column: 20

Seq primer: CCTATTAGGTGACACTATAGAAC

High quality sequence stop: 585.

FEATURES

source Location/Qualifiers

1..585

/organism="Bos taurus"

/mol\_type="mRNA"

/strain="Holstein"

/db\_xref="taxon:9913"

/clone="8BOV\_51K20"

/sex="Female"

/tissue\_type="Epithelial, Muscle"

/dev\_stage="Lactating, Neonatal"

/lab\_host="DH10B TON"

/clone\_lib="BARC 8BOV"

/notes="Organ: Intestine; Vector: pCMVSPORT6.1; Site: 1: Noti; Site 2: EORI; Normalized cow cDNA intestinal library in pCMVSPORT6.1, constructed from equimolar mRNA pools derived from 5 sources, 4 lactating intestinal, 1 neonatal intestinal 4/5 Lactating, Proximal Duodenum, Jejunum, Distal Ileum, Colon, 1/5 Neonatal, Proximal Duodenum, Jejunum, Distal Ileum"

ORIGIN

Alignment Scores:

Pred. No.: 2,12e-51 Length: 585

Score: 940.00 Matches: 192

Percent Similarity: 98.97% Conservative: 0

Best Local Similarity: 98.97% Mismatches: 2

Query Match: 49.21% Indels: 0

DB: 7 Gaps: 0

US-09-544-776-2 (1-373) x CO259245 (1-585)

Qy 176 ProLysArgArgGlySerValValValAspLeuLeuTyrTrpArgAspIle 195

Db 2 CCCAAACGCGGGGCTCCTCGGGCTCAGTGGTGTGGTACCTCTCTACTGGAGACATT 61

Qy 196 LysLysThrGlyValValPheGlyAlaSerLeuPheLeuLeuSerLeuThrValPhe 215

Db 62 AAGAAGACTGGAGTGGTGGTCCGGTGCAGCTGTTCCTGCTGCTCGCTGACAGTATTC 121

Qy 216 SerIleValSerValThrAlaTyrIleAlaLeuAlaLeuSerValThrIleSerPro 235

Db 122 AGCATTTGTGAGTGAACGGCTACATTGGCTTGGCCCTGCTCTCTGTGACTATCAGCTTT 181

Qy 236 ArgIleTyrIysGlyValIleGlnAlaIleGlnLysSerAspGluGlyHisProPheArg 255

Db 182 AGGATATATAAGGGGTGTGATCCAGGCTATCCAGAAATCTGATGAAGGCCACCCATTGAG 241

Qy 256 AlaTyrLeuGluSerGluValAlaIleSerGluGluLeuValGlnLysTyrSerAsnSer 275

Db 242 GCATATTTGGATCTGAAGTTGCTATATCTGAGGAGTTGGTTCAGAAAGTACAGCAATTCT 301

Qy 276 AlaLeuGlyHisValAenCysThrIleLysGluLeuArgArgLeuPheLeuValAspAep 295

Db 302 GCTCTTTGGTCACTGTTAACTGCACATATAAAGAACTCAGACGCCCTCTCTTAGTTGATGAT 361

Qy 296 LeuValAspSerLeuLysPheAlaValLeuMetTrpValPheThrTyrValGlyAlaLeu 315

Db 362 TTAGTTTGGATTTCTCTGAAGTTTGCAGTGTGATGTTGGGTATTTTACCTATGTTGGTGGCTG 421

Qy 316 PheAsnGlyLeuThrLeuLeuLeuAlaLeuLysSerLeuPheSerValProValIle 335

Db 422 TTCATGTTGTTGACACTACTAATTTTGGCTCTGATTTTCACTCTCTCAGTGTCTCTGTTATT 481

Qy 336 TyrGluArgHisGlnAlaGlnIleAspHisTyrLeuGlyLeuAlaAenLysAenValLys 355

Db 482 TATGAACGGCATCAGCGCAATATAGATCATTTATCTGGAGCTTCGCAATTAAGATGTTAAA 541

Qy 356 AspAlaMetAlaLysIleGlnAlaLysIleProGlyLeuLys 369

Db 542 GATGCTATGGCTAAATATCCAGCAAAATCCCTGGATTGAAG 583

RESULT 14

LOCUS BF099705

DEFINITION 601751340F1 NCI\_CGAP\_Mam1 Mus musculus cDNA clone IMAGE:3979233 5', mRNA sequence.

ACCESSION BF099705

VERSION BF099705.1 GI:10882231

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 815)

AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgabs-r@mail.nih.gov](mailto:cgabs-r@mail.nih.gov)  
Tissue Procurement: Gilbert Smith, Ph.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Plate: LLAM9172 row: 0 column: 10  
High quality sequence stop: 712.

FEATURES

source Location/Qualifiers

1..815



QY	245	IleGlnLysSerAspGluGlyHisProPheArgAlaTyrLeuGluSerGluValAlaIle	264
Db	243	ATCCAGAAATCAGATGAAGCCACCCATTCTGAGGCATATCTGGAATCTGAAGTTGCTATA	302
QY	265	SerGluGluLeuValGlnLysTyrSerAsnSerAlaLeuGlyHisValAsnCysThrIle	284
Db	303	TCTGAGGAGTTGGTTCAGAAAGTACAGTAAATCTGCTCTTGSTCATGTGAACGACGATA	362
QY	285	LysGluLeuArgLeuPheLeuValAspLeuValAspSerLeuLysPheAlaVal	304
Db	363	AAGAACTCAGGCGCCTCTTTAGTTGATGATTTAGTTGATTTCTGAAAGTTTGCAATG	422
QY	305	LeuMetTrpValPheThrTyrValGlyAlaLeuPheAsnGlyLeuThrLeuLeuIleLeu	324
Db	423	TTGATGTGGGTATTTACCTATGTTGGTGCCTTGTGTTAATGCTGACACTACTGATTTTG	482
QY	325	AlaLeuIleSerLeuPheSerValProValIleTyrGluArgHisGlnAlaGlnIleAsp	344
Db	483	GCTCTCATTTTCACTCTTCACTGTTCCCTGTTATTTATGAAACGGCATCAGGCACAGATAGAT	542
QY	345	HisTyrLeuGlyLeuAlaAsnLysAsnValLysAspAlaMetAlaLysIleGlnAlaLys	364
Db	543	CATTATCTACGACTTGCAAAATAGAAATGCTAAAGATGCTATGGCTAAATCCAAGCAAAA	602
QY	365	IleProGlyLeuLysArgLysAlaGlu	373
Db	603	ATCCCTGGATTGAAGCCCAAGCTGAA	629

Search completed: August 4, 2005, 00:29:15  
Job time : 3113 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 3, 2005, 17:13:52 ; Search time 207 Seconds  
(without alignments)

2948.460 Million cell updates/sec

Title: US-09-544-776-2

Perfect score: 1910

Sequence: 1 MEDDQSPVSSSDSPRPQ.....VKDAMAKIQAKIPGLKRRAE 373

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp  
-Q=/cgn2\_1/USPTO\_spool\_p/US09544776/runat\_03082005\_181346\_28231/app\_query.fasta\_1.519  
-DB=Issued Patents NA -Qfmt=fastap -SUFFIX=p2n.rn1 -MINMATCH=0.1 -LOOPEL=0  
-LOOPEXT=0 -UNITS=Bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=us09544776 @CGN 1.1.105 @runat\_03082005\_181346\_28231 -NCPU=6 -ICPU=3  
-NO\_WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA:\*  
1: /cgn2\_6/ptodata/1/ina/5A COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PCTUS COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1431	74.9	4822	3	US-09-484-970B-106
2	917	48.0	799	2	Sequence 106, App
3	917	48.0	1669	4	Sequence 2, Appli
4	904	47.3	2610	4	Sequence 3253, Ap
5	747.5	39.1	2069	4	Sequence 382, App
6	747.5	39.1	3202	4	Sequence 3309, Ap
7	672.5	35.2	1766	3	Sequence 1127, Ap
8	659.5	34.5	2664	3	Sequence 254, App
9	613.5	32.1	2662	4	Sequence 255, App
10	550	28.8	1095	2	Sequence 2988, Ap
11	547	28.6	3517	4	Sequence 4, Appli
12	541.5	28.4	794	3	Sequence 111, App
					Sequence 102, App

13	527.5	27.6	2181	4	US-09-949-016-1419	Sequence 1419, Ap
14	513	26.9	2014	4	US-09-270-767-13561	Sequence 13561, A
15	347	18.2	441	4	US-09-513-999C-2227	Sequence 2227, Ap
16	344.5	18.0	454	4	US-09-621-976-740	Sequence 740, App
17	344.5	18.0	463	4	US-09-621-976-741	Sequence 741, App
c 18	332	17.4	601	4	US-09-949-016-117588	Sequence 117588,
c 19	332	17.4	601	4	US-09-949-016-117589	Sequence 117589,
c 20	332	17.4	42075	4	US-09-949-016-14995	Sequence 14995, A
21	301	15.8	443	4	US-09-513-999C-3784	Sequence 3784, Ap
22	275	14.4	261	2	US-08-700-607-9	Sequence 9, Appli
23	269.5	14.1	135667	4	US-09-949-016-15051	Sequence 15051, A
24	269.5	14.1	152486	4	US-09-949-016-12869	Sequence 12869, A
c 25	266	13.9	601	4	US-09-949-016-10169	Sequence 10169, A
c 26	266	13.9	601	4	US-09-949-016-119335	Sequence 119335,
c 27	257	13.5	13906	4	US-09-949-016-14730	Sequence 14730, A
28	256	13.4	200	4	US-09-513-999C-11198	Sequence 11198, A
29	246	12.9	15661	4	US-09-949-016-13161	Sequence 13161, A
c 30	217	11.4	2580	3	US-09-050-863-2	Sequence 2, Appli
c 31	217	11.4	2580	3	US-09-359-081-2	Sequence 2, Appli
c 32	217	11.4	8705	4	US-09-647-344A-14	Sequence 14, Appli
c 33	217	11.4	9600	3	US-08-910-647-1	Sequence 1, Appli
c 34	217	11.4	9600	3	US-09-620-925-1	Sequence 1, Appli
c 35	217	11.4	10596	1	US-07-884-811-15	Sequence 15, Appl
c 36	217	11.4	10596	1	US-07-885-971-15	Sequence 15, Appl
c 37	217	11.4	10596	1	US-08-087-783A-15	Sequence 15, Appl
c 38	217	11.4	10596	1	US-08-194-088B-15	Sequence 15, Appl
c 39	217	11.4	10596	1	US-08-194-087-15	Sequence 15, Appl
c 40	217	11.4	10596	5	PCT-US93-04648-15	Sequence 15, Appl
41	217	11.4	16080	4	US-09-724-566A-48	Sequence 48, Appl
42	217	11.4	16080	4	US-09-471-669A-48	Sequence 48, Appl
43	212	11.1	5452	2	US-09-130-114-1	Sequence 1, Appli
c 44	210	11.0	1926	3	US-09-249-585A-2	Sequence 2, Appli
c 45	210	11.0	1926	4	US-09-410-399-3	Sequence 3, Appli

#### ALIGNMENTS

##### RESULT 1

US-09-484-970B-106  
; Sequence 106, Application US/09484970B  
; Patent No. 6426186  
; GENERAL INFORMATION:  
; APPLICANT: Jones, Karen A.  
; APPLICANT: Volkmut, Wayne  
; APPLICANT: Walker, Michael G.  
; TITLE OF INVENTION: BONE REMODELING GENES  
; FILE REFERENCE: PB-0014 US  
; CURRENT APPLICATION NUMBER: US/09/484,970B  
; CURRENT FILING DATE: 2000-01-18  
; NUMBER OF SEQ ID NOS: 172  
; SOFTWARE: PERL Program  
; SEQ ID NO 106  
; LENGTH: 4822  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. 6426186 444857.15CBI  
; NAME/KEY: unsure  
; LOCATION: 33, 51, 79, 211, 369, 483-484, 731, 748, 4803, 4805-4806, 4808-4809,  
; OTHER INFORMATION: a, t, c, g, or other  
US-09-484-970B-106

Alignment Scores:  
Pred. No.: 2.69e-89  
Score: 1431.00  
Percent Similarity: 30.68%  
Best Local Similarity: 30.60%  
Query Match: 74.92%  
DB: 3  
Length: 4822  
Matches: 365  
Conservative: 1  
Mismatch: 7  
Indels: 820  
Gaps: 1  
US-09-544-776-2 (1-373) x US-09-484-970B-106 (1-4822)

QY	1	MetGluAspLeuAspGlnSerProLeuValSerSerSerAspSerProArgProGln	20
Db	253	ATGGAAGACTGGACGAGTCTCTCTGGTCTCGTCTCGACAGCCACCCCGCGCGAG	312
QY	21	ProAlaPheLysTyrglnPheValArgGluProGluAspGluGluGluGluGlu	40
Db	313	CCCGGCTTCAAGTACCAGTTCTGTGGAGGAGCCGAGGACGAGGAGGAAGAGGANGAT	372
QY	41	GluGluGluAspGluAspGluAspLeuGluLeuGluValLeuGluArgLysProAla	60
Db	373	GAAGAGGAGGACGAGGACCAAGACCTTGGAGGAGCTGGAGGTCTGGAGAGGAGCCCGCC	432
QY	61	AlaGlyLeuSerAlaAlaProValProThrAlaProAlaAlaGlyValaProLeuMetAsp	80
Db	433	GCGGGCTGTCCGGGCCCCAGTGCACCGCCCTGCGCCGCGCGGCNNNTAATGGAC	492
QY	81	PheGlyAsnAspPheValProProAlaProAlaProArgGlyPheLeuProAlaAlaProProVal	100
Db	493	TTTCGGAATGACTTCGTGCGCGCGCGCCCGGGACCCCTGCGGCGCTCCCGCCGTC	552
QY	101	AlaProGluArgGlnProSerTrpAspProSerProValSerSerThrValProAlaPro	120
Db	553	GCCCCGAGCGCAGCGCTTGGGACCGCGCGGTGCTGCGACCGTGCCTCCCGCGCA	612
QY	121	SerPheLeuSerAlaAlaValSerProSerLysLeuProGluAspGluProPro	140
Db	613	TCCCCGTGTCTGTGCGCCAGTCTCGCCCTCCAGCTCCCTGAGGACGACGAGCTCCG	672
QY	141	AlaArgProProProProProAlaSerValSerProGlnAlaGluProValTrpThr	160
Db	673	GCCCCGCTCCCTCCTCCCGCGCAGCGTGAGCCCGCCAGGAGCCCGTGTGGANC	732
QY	161	ProProAlaProAlaProAlaProSerThrProAlaProLysArgArgGly	180
Db	733	CGCGCAGCCCCGGCTNCCGCGCGCCCCCTCCACCCGCGCGCCCAAGCGCAGGGGC	792
QY	181	SerSerGlySerVal	185
Db	793	TCTCGGGCTCAGTGATGAGACCCCTTTTGTGCTCTCTCTGCTGATCTGAGCCTGTGATA	852
QY	185	-----	185
Db	853	CGCTCCTCTGCAGAAAAATAGCATGAAGGAGCAGCAGGTAAACACTATTTTCGGCTGGT	912
QY	185	-----	185
Db	913	CAAGAGATTCCCATTCTCTCTGTGTTGAACTGCTGTTCTCTCTCTCTCTCTCTCTCT	972
QY	185	-----	185
Db	973	CTCTCAGCCGCTCTTTCAGAGAACATGAATACCTTGGTAATTTGTCAACAGTATTACCC	1032
QY	185	-----	185
Db	1033	ACTGAAGGAACACTTCAAGAAAATGTCACTGAAGCTTCTAAAGAGCTCTCAGAGAAGCA	1092
QY	185	-----	185
Db	1093	AAAACTCTACTCATAGATAGATTAAACAGAGTTTTCAGATTAGATTAGATTAGATTAG	1152
QY	185	-----	185
Db	1153	GGATCATCGTTCAGTGTCTCTCCAAAAGCAGAAATCTGCCGTAAATAGTACCAATCCTAGG	1212
QY	185	-----	185
Db	1213	GAAGAAATAATCGTGAAAAATAAAGATGAAGAGAGAGTGTAGTAGTAATACATCCTT	1272
QY	185	-----	185
Db	1273	CATAATCAACAGAGTTTACCTACAGCTCTTACTAAATGGTTAAAGAGGATGAAGTTTG	1332
QY	185	-----	185
Db	1333	TCCTTCAGAAAAAGCAAAAGACAGTTTTAAATGAAAGAGAGTTGCAGTGGAGCTCCTATG	1392
QY	185	-----	185
Db	1393	AGGGAGGAATATGCAGACTTCAAAACATTTGAGCGAGTATGGGAAGTGAAGATAGTAAG	1452
QY	185	-----	185
Db	1453	GAAGATAGTATATGTTGGCTGCTGGAGTAAATCGAGAGCAACTTGGAAAAGTAAAGTG	1512
QY	185	-----	185
Db	1513	GATAAAAATGTTTTCAGATAGCCTTGCAGCAAACTAATCACGAAAAAGATAGTAGACT	1572
QY	185	-----	185
Db	1573	AGTAATGATGATACCTTCTTTCCAGTAGCCGAGGAGGTATAAGGATCGTTTCAGAGCA	1632
QY	185	-----	185
Db	1633	TATATCATGTGTCTCCCTTTAAACCAGAGCAACTGAGAGCATTTGCAACAAACATTTT	1692
QY	185	-----	185
Db	1693	CTTTTGTAGGAGATCTTACTTCAGAAAATAAGACCGATGAAAAAATAAGAAAA	1752
QY	185	-----	185
Db	1753	GAAGCCCCAAATAGTAACAGAGAGAATACTAGCACCAAAACATCAAAACCTTTTACTTG	1812
QY	185	-----	185
Db	1813	TAGCAGCAGAGTTCTCGAGACAGATTATGTCAACAACAGATTAATTAACAAGGTGACTG	1872
QY	185	-----	185
Db	1873	AGGAAGTCGTGCAAAACATGCCTGAAGCCTGACTCCAGATTTAGTACAGGAAGCATGTG	1932
QY	185	-----	185
Db	1933	AAAGTGAATTGAATGAAGTTACTGGTACAAGATTGCTTATGAACAACAAAAATGGACTGG	1992
QY	185	-----	185
Db	1993	TTCAAAACATCAGAGTTATGCAAGAGTCACTCTATCTCTGCAGCAGAGCTTTGCCCATCAT	2052
QY	185	-----	185
Db	2053	TTGAAGAGTCAGAGCTACTCTTCACCAGTTTTCCTGCTGACATTGTATGGAAGCACCAT	2112
QY	185	-----	185
Db	2113	TGAATTCCTGCAGTTCCTAGTGTGTGCTTCGTGATACAGCCAGCTCATCCCATTAG	2172
QY	185	-----	185
Db	2173	AAGCTTCTTCAGTTAATTAAGACATAAACAATGAGCCTGAAACCCCCACCATATG	2232
QY	185	-----	185
Db	2233	AAGAGCCATGAGTGTATCACTAAAAAAGTATCAGGAATAAAGGAAGAAATTAAGAGC	2292
QY	185	-----	185
Db	2293	CTGAAAAATTAATTCAGCTCTTCAAGAAACAGAGCTCTTATATATATCTATTGCAATG	2352
QY	185	-----	185
Db	2353	ATTTAATTAAGAAAAAAGCTTTCTGCTGAACAGCTCCGATTTCTCTGATTATTTCAG	2412
QY	185	-----	185



Db 2413 AATGGCAAAAGTTGAACAGCCAGTGCCTGATCATTTCTGAGCTAGTTGAAGATTCCTCAC 2472  
QY 185 ----- 185  
Db 2473 CTGATTTGNAACCAAGTTGACTTATTAGTAGTATGATGATTCATACCTGAGCTTCCACAAAAC 2532  
QY 185 ----- 185  
Db 2533 AAGATGAACACTGTGATGCTTTGTGAAGAAAGTCTCACTGAGACTTTCATTTGAGTCAATGA 2592  
QY 185 ----- 185  
Db 2593 TAGAATATGAATAAAGGAAAACTCAGTGTCTTTGGCCACTGAGGGAGAAAGCCATATT 2652  
QY 185 ----- 185  
Db 2653 TGGAAATCTTTTAAGCTCAGTTTAGATAACACAAAAGNATCCCTGTGTACCTGATGAAGTTT 2712  
QY 185 ----- 185  
Db 2713 CAACATTTGAGCAAAAGGAAAAATTCCTTTGCAGATGGAGGAGCTCAGTACTGCGATTT 2772  
QY 185 ----- 185  
Db 2773 ATTCAATGATGACTTATTATTCTAAGGAAGCACAGATAAGAGAACTGAACGTTTTT 2832  
QY 185 ----- 185  
Db 2833 CAGATTCATCTCAAATGAAATATATAGATGAGTTCCCTACATGTATGATCAGTTCTTAAACTG 2892  
QY 185 ----- 185  
Db 2893 ATTCATTTCTAAATTAGCCAGGGAATATATCTGACCTAGAAGTATCCACAAAAGTGAAA 2952  
QY 185 ----- 185  
Db 2953 TTGCTAATGCCCGGATGGAGCTGGGTCAATGCTTGCACAGAAATGGCCCCATGACCTTT 3012  
QY 185 ----- 185  
Db 3013 CTTTGAAGAACATACAAACCAAGTTGAAGAGAAATCAGTTTCTCAGATGACTTTTCTA 3072  
QY 185 ----- 185  
Db 3073 AAAATGGGTCTGTACATCAAGGTGCTCTTATTGCTCCAGATGTTTCTGCTTTGGCCA 3132  
QY 185 ----- 185  
Db 3133 CTCAGCAGAGATAGAGAGCATAGTTAAACCCAAAGTCTTGTGAAGAAGCTGAGAAA 3192  
QY 185 ----- 185  
Db 3193 AACTTCCTCCGATACAGAAAAGAGGACAGATCACCATCTGTATATTTTCAGCAGAGC 3252  
QY 186 -----ValValAspLeuLeuTyrTrpArgAspIleLysLysThrGlyVal 200  
Db 3253 TGAGCTAAACCTTCAGTTGTGACCTCTGTACTGAGAGACATTAAGAAGACTGGAGTG 3312  
QY 201 ValPheGlyValAspLeuPheLeuLeuLeuSerLeuThrValPheSerIleValSerVal 220  
Db 3313 GTGTTTGGTGCCAGCTATTCTGCTGCTTTCAATGCACAGTATTCAGCATTTGAGCGTA 3372  
QY 221 ThrAlaTyrIleAlaLeuAlaLeuSerValThrIleSerProArgIleTyrLysGly 240  
Db 3373 ACAGCTCATATGCTTGGCCCTGCTCTCTGTGACCATCAGCTTTTAGGATATACAGGGT 3432  
QY 241 ValIleGlnAlaIleGlnLysSerAspGluGlyHisProPheArgAlaTyrLeuGluSer 260  
Db 3433 GTGATCCAAGCTATCCAGAAATCAGATGAAGGCCACCCATTCAGGCGCATATCTGGAATCT 3492  
QY 261 GluValAlaIleSerGluGluLeuValGlnLysTyrSerAsnSerAlaLeuGlyHisVal 280  
Db 3493 GAAGTTGCTATATCTGAGGAGTTGGTTTGAAGAGTACAGTAATTTCTGCTCTTGGTCAATG 3552

QY 281 AsnCysThrIleLysGluLeuArgArgLeuPheLeuValAspAspLeuValAspSerLeu 300  
Db 3553 AACTGCACGATAAAGGAACCTCAGGCGCTCTCTCTTAGTTGATGATTTAGTTGTTCTCTG 3612  
QY 301 LysPheAlaValLeuMetTrpValPheThrTyrValGlyAlaLeuPheAsnGlyLeuThr 320  
Db 3613 AAGTTTGCAGTGTGATGTGGGTATTACCTATGTTGGTGCCTGTTTAATGGTCTGACA 3672  
QY 321 LeuLeuIleAlaLeuIleSerLeuPheSerValProValIleTyrGluArgHisGln 340  
Db 3673 CTACTGATTTGGCTCTCAATTCACCTTCAGTGTTCCTGTTATTATTATGACGGCATCAG 3732  
QY 341 AlaGlnIleAspHisTyrLeuGlyLeuAlaAsnLysAsnValLysAspAlaMetAlaLys 360  
Db 3733 GCACAGATAGATCATTTATCTAGGACTTGCAATAAGAAATGTTAAAGATGCTATGGCTAAA 3792  
QY 361 IleGlnAlaLysIleProGlyLeuLysArgLysAlaGlu 373  
Db 3793 ATCCAAGCAAAAATCCCTGGATTGAAGCGCAAGCTGAA 3831

## RESULT 2

US-08-700-607-2  
; Sequence 2, Application US/08700607  
; Patent No. 5858708  
; GENERAL INFORMATION:  
; APPLICANT: Bandman, Olga  
; APPLICANT: Au-Young, Janice  
; APPLICANT: Goli, Surya K.  
; APPLICANT: Hillman, Jennifer L.  
; TITLE OF INVENTION: TWO NOVEL HUMAN NSP-LIKE PROTEINS  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: U.S.  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 1.5  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/700,607  
; FILING DATE: Filed Herewith  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0114 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 799 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; IMMEDIATE SOURCE:  
; LIBRARY:  
; CLONE: Consensus  
; US-08-700-607-2

Alignment Scores:  
Pred. No.: 8,12e-55 Length: 799  
Score: 917.00 Matches: 187  
Percent Similarity: 99.47% Conservative: 0  
Best Local Similarity: 99.47% Mismatches: 1  
Query Match: 48.01% Indels: 0  
DB: 2 Gaps: 0

US-09-544-776-2 (1-373) x US-08-700-607-2 (1-799)

```
QY 186 ValValAspLeuLeuTyrTrpArgAspIleLysLysThrGlyValValPheGlyAlaSer 205
Db 108 GTTGTGACCTCTCTGACTGAGAGACATTAAGAAGACTGGAGTGGTGTGGTGCACG 167
QY 206 LeuPheLeuLeuSerLeuThrValPheSerIleValSerValThrAlaTyrIleAla 225
Db 168 CTATTCTCTGCTCTCTGACTGAGTATTCAGCATTTGTGAGCGTAAACGCCATACATTGCC 227
QY 226 LeuAlaLeuLeuSerValThrIleSerProArgIleTyrLysGlyValIleGlnAlaIle 245
Db 228 TTGGCCCTCTCTCTGTGACCATCAGCTTTAGGATATACAAGGGTGTGATCCAAGCTATC 287
QY 246 GlnLysSerAspGluGlyHisProPheArgAlaTyrLeuGluSerGluValAlaIleSer 265
Db 288 CAGAAATCAGATGAAGGCCACCANTTCAGGGCATATCTGGAATCTGAACTTCTATATCT 347
QY 266 GluGluLeuValGlnLysTyrSerAsnSerAlaLeuGlyHisValAsnCysThrIleLys 285
Db 348 GAGGAGTTGGTTCAGAGTACAGTAATCTCTCTTTGGTCATGTGAATGCGACGATAAAG 407
QY 286 GluLeuArgArgLeuPheLeuValAspLeuValAspSerLeuLysPheAlaValLeu 305
Db 408 GAACTCAGCGCCTCTCTTAGTTGATGATTAGTTGATCTCTGAAGTTTCAGGTGTG 467
QY 306 MetTrpValPheThrTyrValGlyAlaLeuPheAsnGlyLeuThrIleLeuAla 325
Db 468 ATGTGGGTATTACCTATGTTGGTCTGTTTAAATGGTCTGACACTACTGATTTTGGCT 527
QY 326 LeuIleSerLeuPheSerValProValIleTyrGluArgHisGlnAlaGlnIleAspHis 345
Db 528 CTCATTTCACTCTTCAGTCTCTGTTATTATGAACGGCATCAGGCACAGATAGATCAT 587
QY 346 TyrLeuGlyLeuAlaAsnLysAsnValLysAspAlaMetAlaLysIleGlnAlaLysIle 365
Db 588 TATCTAGGACTTGCAATAAGAATGTTAAAGATGCTATGGCTAAATCCAGCAAAAATC 647
QY 366 ProGlyLeuLysArgLysAlaGlu 373
Db 648 CTTGGATTGAAGCGCAAAAGCTGAA 671
```

## RESULT 3

```
US-09-949-016-3253
; Sequence 3253, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3253
; LENGTH: 1669
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-3253
```

Alignment Scores: 2,148-54 Length: 1669  
Pred. No.: 917.00 Matches: 187  
Score: 99.47% Conservative: 0  
Percent Similarity: 99.47% Mismatches: 1  
Best Local Similarity: 99.47%

Query Match: 48.01% Indels: 0  
DB: 4 Gaps: 0

US-09-544-776-2 (1-373) x US-09-949-016-3253 (1-1669)

```
QY 186 ValValAspLeuLeuTyrTrpArgAspIleLysLysThrGlyValValPheGlyAlaSer 205
Db 248 GTTGTGACCTCTCTGACTGAGAGACATTAAGAAGACTGGAGTGGTGTGGTGCACG 307
QY 206 LeuPheLeuLeuSerLeuThrValPheSerIleValSerValThrAlaTyrIleAla 225
Db 308 CTATTCTCTGCTCTCTGACTGAGTATTCAGCATTTGTGAGCGTAAACGCCATACATTGCC 367
QY 226 LeuAlaLeuLeuSerValThrIleSerProArgIleTyrLysGlyValIleGlnAlaIle 245
Db 368 TTGGCCCTCTCTCTGTGACCATCAGCTTTAGGATATACAAGGGTGTGATCCAAGCTATC 427
QY 246 GlnLysSerAspGluGlyHisProPheArgAlaTyrLeuGluSerGluValAlaIleSer 265
Db 428 CAGAAATCAGATGAAGGCCACCANTTCAGGGCATATCTGGAATCTGAACTTCTATATCT 487
QY 266 GluGluLeuValGlnLysTyrSerAsnSerAlaLeuGlyHisValAsnCysThrIleLys 285
Db 488 GAGGAGTTGGTTCAGAGTACAGTAATCTCTCTTTGGTCATGTGAATGCGACGATAAAG 547
QY 286 GluLeuArgArgLeuPheLeuValAspLeuValAspSerLeuLysPheAlaValLeu 305
Db 548 GAACTCAGCGCCTCTCTTAGTTGATGATTAGTTGATCTCTGAAGTTTCAGGTGTG 607
QY 306 MetTrpValPheThrTyrValGlyAlaLeuPheAsnGlyLeuThrLeuLeuAla 325
Db 608 ATGTGGGTATTACCTATGTTGGTCTGTTTAAATGGTCTGACACTACTGATTTTGGCT 667
QY 326 LeuIleSerLeuPheSerValProValIleTyrGluArgHisGlnAlaGlnIleAspHis 345
Db 668 CTCATTTCACTCTTCAGTCTCTGTTATTATGAACGGCATCAGGCACAGATAGATCAT 727
QY 346 TyrLeuGlyLeuAlaAsnLysAsnValLysAspAlaMetAlaLysIleGlnAlaLysIle 365
Db 728 TATCTAGGACTTGCAATAAGAATGTTAAAGATGCTATGGCTAAATCCAGCAAAAATC 787
QY 366 ProGlyLeuLysArgLysAlaGlu 373
Db 788 CTTGGATTGAAGCGCAAAAGCTGAA 811
```

## RESULT 4

```
US-09-023-655-382
; Sequence 382, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESS: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/023,655
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
```

```
/ APPLICATION NUMBER:
/ FILING DATE:
/ CLASSIFICATION:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Zeller, Karen J.
/ REGISTRATION NUMBER: 37,071
/ REFERENCE/DOCKET NUMBER: PA-0001 US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (650) 855-0555
/ TELEFAX: (650) 845-4166
/ INFORMATION FOR SEQ ID NO: 382:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 2610 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ IMMEDIATE SOURCE:
/ LIBRARY: LUNGNOT14
/ CLONE: 1508778
/ US-09-023-655-382

Alignment Scores:
Pred. No.: 3,04e-53 Length: 2610
Score: 504.00 Matches: 187
Percent Similarity: 98.94% Conservative: 0
Best Local Similarity: 98.94% Mismatches: 1
Query Match: 47.33% Indels: 1
DB: 4 Gaps: 0

US-09-544-776-2 (1-373) x US-09-023-655-382 (1-2610)

QY 186 ValValAspLeuLeuTyrTrpArgAspIleLysLysThrGlyValValPheGlyAlaSer 205
DB 1311 GTTGTGACCTCTGTACTGAGAGACATTAAAGAGACTGGAGTGGTGTGGTGCAGC 1370
QY 206 LeuPheLeuLeuLeuSerLeuThrValPheSerIleValSerValThrAlaTyr-IleAl 225
DB 1371 CTATTCTCTGCTGCTTCAATTGACAGATATTTCAGCATTTGTGAGCGTAAACAGCTACAATTGC 1430
QY 225 aLeuAlaLeuLeuSerValThrIleSerProArgIleTyrLysGlyValIleGlnAlaI 245
DB 1431 CTTGGGCCCTCTCTCTGTGACCATCAGCTTTAGGATATACAAAGGTGTGATCCAGCTAT 1490
QY 245 eGlnLysSerAspGluGlyHisProPheArgAlaTyrLeuGluSerGluValAlaIleSe 265
DB 1491 CCAGAAATCAGATGAAGGCCACCCATTTCAGGCATATCTGGATCTGAAGTTCGTATATC 1550
QY 265 rGluGluLeuValGlnLysTyrSerAsnSerAlaLeuGlyHisValAsnCysThrIleLy 285
DB 1551 TGAGGAGTTGGTTTCAGAAATGACAGTAAATCTCTGCTTGTGTCATGTGAACCTGCACGATAA 1610
QY 285 sGluLeuArgArgLeuPheLeuValAspLeuValAspLeuValAspLeuLysPheAlaVal 305
DB 1611 GGAACCTCAGCGGCTCTCTTAGTGTATGATTTAGTGTATCTCTGAAGTTTCAGTGT 1670
QY 305 uMetTrpValPheThrTyrValGlyAlaLeuPheAsnGlyLeuThrLeuLeuIleLeuAl 325
DB 1671 GATGTGGGTATTTACCTATGTTGGTCTGTTGTTTAAATGTTCTGACACTACTGATTTTGGC 1730
QY 325 aLeuIleSerLeuPheSerValProValIleTyrGluArgHisGlnAlaGlnIleAspHi 345
DB 1731 TCTCATTTCACTTTCAGTGTCTCTGTTATTTATTAACGGCATCAGGCACAGATAGATCA 1790
QY 345 sTyrLeuGlyLeuAlaAsnLysAsnValLysAspAlaMetAlaLysIleGlnAlaLysI 365
DB 1791 TTAATCAGGCTTGCAATTAAGAAATGTTAAAGATGCTATGGCTAAATAATCCAGCAAAAT 1850
QY 365 eProGlyLeuLysArgLysAlaGlu 373
DB 1851 CCCTGGGTGAAGCGCAAAAGCTGAA 1875

RESULT 5
US-09-949-016-3309

; Sequence 3309, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3309
; LENGTH: 2069
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-3309

Alignment Scores:
Pred. No.: 1.45e-42 Length: 2069
Score: 747.50 Matches: 180
Percent Similarity: 54.12% Conservative: 50
Best Local Similarity: 42.35% Mismatches: 113
Query Match: 39.14% Indels: 83
DB: 4 Gaps: 10

US-09-544-776-2 (1-373) x US-09-949-016-3309 (1-2069)

QY 4 LeuAspGlnSerProLeuValSerSerSerAspSerProProArgProGlnProAlaPhe 23
DB 127 CTGCGCCAGACAGCCCGAGGTCGAAGGCCAGGTCCGCGACCGCCCAACC----- 171
QY 24 LysTyrGlnPheValArgGluProGluAspGluGluGluGluGluGluGluGluGlu 43
DB 172 -----ATCCCCAGCCCCCTCGACACAGCGCCAGCGCGGAGTCCGGGGAC 219
QY 44 AspGluAspGluAspLeuGluLeuGluValLeuGluArgLysProAlaAlaGlyLeu 63
DB 220 -----TCAGAGATCGAGCTGTGTCCGAGGACCCCATTCGCGCGGAG 261
QY 64 SerAlaAlaProValProThrAlaProAlaAlaGlyAlaProLeuMetAspPheGlyAsn 83
DB 262 GAGCGCTGCCCC-----TCAGGCTATGTGAGCTTTGGCCAC 297
QY 84 AspPheValProProAlaProArgGlyPheLeuProAlaAlaProProVal----- 100
DB 298 GTGGCGCGCCCGCGCCCTCG-----CCGCGCTCGCCATCCATCCAGTACAGC 345
QY 101 -----AlaProGluArgGln-ProSerTrpAspProSerProValSerSerThrVal-- 117
DB 346 ATCCTGAGGAGGAGCGCGCGCGAGCTGGACAGCGAGCTCATCATCGAGTCTGTCGAC 405
QY 118 -----Pr 118
DB 406 GCCTCTCGGCTCGGAGGAGAGCCCAAGCGGAGCAGACTCACCCCGATGAAGCCC 465
QY 118 oAlaPro-----SerPheLeuSerAl 125
DB 466 AGCGCCCTGGATGCCATCCGGGAGGAGACTGGCGTCCGGGCCGAGGAGCGTGCGCCAAGC 525
QY 125 aAlaAlaValSerProSerLysLeuProGluAspAspGluProProAlaArgProPr 145
DB 526 CGGCGGGGCTCGCGAGCGCGGTTCCTCTCCTC-GACTACCCCTCACTGAGGCCCGCAGCC 584
QY 145 oPro-----ProProAlaSerValSerProGlnAlaGluProValTrpThrProPr 162
DB 585 TGGCCCCGAGTGGCCCCCTGGAGAGCGGCTGGAGCGCTTGAGAGCCCTGTTGCCACG 644
```

```
QY 162 oAlaPro---AlaProAlaAlaProSerThrProAlaAlaProLysArgGlySe 181
Db 645 GAAGCTGAGAGACTCGAGTCCAAACCAAGTCTCGGGCCACAAAGGGCCCTGGGCC 704
QY 181 rSerGly-----SerValValValAspLeu 189
Db 705 TCTAGTCTCGGGCCGCCGCCACTGCTGTTCTCAATAAGCAAAAGCTATTGACCT 764

QY 189 uLeuTyrTrpArgAspIleLysThrGlyValValPheGlyAlaSerLeuPheLeuLe 209
Db 765 GTTGATTGCGGGACATCAAGCAGCGGCATCGTGTGTGGAGTTTCTCTGCTGCT 824

QY 209 uLeuSerLeuThrValPheSerIleValSerValThrAlaTyrIleAlaLeuAlaLeuLe 229
Db 825 CTTCTCCCTGACCCAGTTCAAGGTGTGAGCTGTGGCTTACCTGGCCCTGGCGCACT 884

QY 229 uSerValThrIleSerProArgIleTyrLysGlyValIleGlnAlaIleGlnLysSerAs 249
Db 885 CTCAGCCACCATCAGTTTCGGCATCTACAAGTCTGTTTACAAAGCAGTGCAGAAAACCGA 944

QY 249 pGluGlyHisProPheArgAlaTyrLeuGluSerGluValAlaIleSerGluGluLeuVa 269
Db 945 CCAAGGCCACCTTTCAAGCCCTACTTGGAGCTTGAGATCACCTTTCTCAGGAGCAGAT 1004

QY 269 lGlnLysTyrSerAsnSerAlaLeuGlyHisValAsnCysThrIleLysGluLeuArgAr 289
Db 1005 TCAGAAGTACAGGACTGCTGCAGTCTTCTACGTGAACAGCACACTTAAGAACTGAGGAG 1064

QY 289 gLeuPheLeuValAspAspLeuValAspSerLeuLysPheAlaValLeuMetTrpValPh 309
Db 1065 GCTCTTCCTGTGCCAGGACTGGTGATTCCTTAAATTTGCAGTCTCTGATGGGCTCCT 1124

QY 309 eThrTyrValGlyAlaLeuPheAsnGlyLeuThrLeuLeuIleLeuAlaLeuIleSerLe 329
Db 1125 GACCTACGTTGGGCTCTCTTCAATGGCTGACCTGCTGCTCATGGCTGTGTTTCAAT 1184

QY 329 uPheSerValProValIleTyrGluArgHisGlnAlaGlnIleAspHisTyrLeuGlyLe 349
Db 1185 GTTACTACTCTGTAGTGTATGTTAAGCACCAGGCACAGATTTGACCAATATCTGGACT 1244

QY 349 uAlaAsnLysAsnValLysAspAlaMetAlaLysIleGlnAlaLysIleProGlyLeuLy 369
Db 1245 TGTGAGGACTCACATAAATGCTGTGTGGCAAGATTTCAGGCTAAATCCAGGGCGCTAA 1304

QY 369 sArgLysAlaGlu 373
Db 1305 GAGGCACGCTGAG 1317

RESULT 6
US-09-949-016-1127
; Sequence 1127, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1127
; LENGTH: 3202
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-1127

Alignment Scores:
Pred. No.: 2,57e-42 Length: 3202
Score: 747.50 Matches: 180
Percent Similarity: 54.12% Conservative: 50
Best Local Similarity: 42.35% Mismatches: 113
Query Match: 39.14% Indels: 83
DB: 4 Gaps: 10

US-09-544-776-2 (1-373) x US-09-949-016-1127 (1-3202)
QY 4 LeuAspGlnSerProLeuValSerSerAspSerProProArgProGlnProAlaPhe 23
Db 1260 CTGGCCGACAGCCCGAGGTCAAGGCCAGGTCCGGACCCCAACC-----1304
QY 24 LysTyrGlnPheValArgGluProGluAspGluGluGluGluGluGluGluGlu 43
Db 1305 -----ATCCCCAGCCCCCTGGACCAAGAGCCGAGCGCGAGTGGGGGGAC 1352
QY 44 AspGluAspGluAspLeuGluGluLeuGluLeuGluLeuGluLeuGluLeu 63
Db 1353 -----TCAGATATCGAGCTGTGTCGAGGACCCCATGGCCGGGAG 1394
QY 64 SerAlaAlaProValProThrAlaProAlaAlaGlyAlaProLeuMetAspPheGlyAsn 83
Db 1395 GACGCGCTGCC-----TCAGGCTATGTGAGCTTTGGCCAC 1430
QY 84 AspPheValProProAlaProArgGlyPheLeuProAlaAlaProProVal-----100
Db 1431 GTGGGCGGCCGCCGCCCTCG-----CCGCGCTCGCCATCCATCCAGTACAGC 1478
QY 101 -----AlaProGluArgGln-ProSerTrpAspProSerProValSerSerThrVal--117
Db 1479 ATCCTGAGGGAGGCGCGGAGCCGAGCTGGACAGGAGCTCATCATCGAGTCGTCGAC 1538
QY 118 -----Pr 118
Db 1539 GCCTCTCGGCTCGAGGAGAGCCCAAGCGGAGCAGAGTCAACCCCGATGAAGCC 1598
QY 118 oAlaPro-----SerPheLeuSerAl 125
Db 1599 AGCGCCCTGGATGTCATCCGGGAGGAGACTGGCGTCCGGGCCGAGAGCGTGGCCCAAGC 1658
QY 125 aAlaAlaValSerProSerLysLeuProGluAspAspGluProProAlaAarProPro 145
Db 1659 CGCGGGGCTGGCCGAGCGGGTTCCTTCCTC-GACTACCCCTCACTAGGCCCGCAGCC 1717
QY 145 oPro-----ProProAlaSerValSerProGlnAlaGluProValTrpThrProPr 162
Db 1718 TGGCCCGGAGCTGCCCTCGGAGAGCGGCCCTGGAGCCTGAGAGCCCATGTTGCCACG 1777
QY 162 oAlaPro---AlaProAlaAlaProProSerThrProAlaAlaProLysArgGlySe 181
Db 1778 GAAGCTGAGAGACTCGAGTTCACAAAGTCTCGGGCCACAAAGGGCCCTGGGCC 1837
QY 181 rSerGly-----SerValValValAspLeu 189
Db 1838 TCTAGTCTCGGGCCGCCGCCACTGCTGTTTCTCAATAAGCAAAAGCTATTGACCT 1897
QY 189 uLeuTyrTrpArgAspIleLysThrGlyValValPheGlyAlaSerLeuPheLeuLe 209
Db 1898 GTTGATTGCGGGACATCAAGCAGCGGCATCGTGTGTGGAGTTTCTCTGCTGCT 1957
QY 209 uLeuSerLeuThrValPheSerIleValSerValThrAlaTyrIleAlaLeuAlaLeuLe 229
Db 1958 CTTCTCCCTGACCCAGTTCAAGGTGTGAGCTGTGGCTTACCTGGCCCTGGCGCACT 2017
QY 229 uSerValThrIleSerProArgIleTyrLysGlyValIleGlnAlaIleGlnLysSerAs 249
Db 2018 CTCAGCCACCATCAGTTTCGGCATCTACAAGTCTGTTTACAAAGCAGTGCAGAAAACCGA 2077
QY 249 pGluGlyHisProPheArgAlaTyrLeuGluSerGluValAlaIleSerGluGluLeuVa 269
Db 2078 CGAAGGCCACCTTTCAAGGGCCTACTTGGAGCTTGAGATCACCTTTCTCAGGAGCAGAT 2137
```

```
QY 269 lGlnLysTyrSerAsnSerAlaLeuGlyHisValAsnCysThrIleLysGluLeuArgAr 289
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2138 TCAGAGTACAGGAGCTGCGCTGAGTCTACGTGAGCAGCACACTTAAGGAACTGAGGAG 2197
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 289 gLeuPheLeuValAspAspLeuValAspSerLeuLysPheAlaValLeuMetTrpValph 309
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2198 GCTCTCTCTGTCAGAGCTGGTGGATTCTTAAATAATTTGCAGTCTGATGTGGCTCCT 2257
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 309 eThrTyrValGlyValAlaLeuPheAsnGlyLeuThrLeuLeuLeuAlaLeuLeuSerLe 329
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2258 GACCTAGTTGGGCTCTCTTCAATGGCCTGACCTGCTGCTCATGGCTGTGGTTCAAT 2317
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 329 uPheSerValProValIleTyrGluArgHisGlnAlaGlnIleAspHisTyrLeuGlyLe 349
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2318 GTTACTCTACTGTAGTGTATGTTAAGCAGCCAGGACACAGATTGACCAATATCTGGGACT 2377
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 349 uAlaAsnLysAsnValLysAspAlaMetAlaLysIleGlnAlaLysIleProGlyLeuLy 369
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2378 TGTGAGGACTCACATAATGCTGTTGTGGCAAGATTTCAGGCTAAATCCAGGCGCTAA 2437
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 369 sArgLysAlaGlu 373
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2438 GAGCAGCGCTGAG 2450
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 7
US-09-149-476-254
; Sequence 254, Application US/09149476
; Patent No. 5420526
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: P2002P1
; CURRENT APPLICATION NUMBER: US/09/149,476
; CURRENT FILING DATE: 1998-09-08
; EARLIER APPLICATION NUMBER: PC/US98/04493
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 60/040,162
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,333
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/038,621
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,626
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,334
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,336
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,163
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/047,600
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,615
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,597
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,502
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,633
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,583
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,617
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,618
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,503
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,592
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,581
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,584
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,500
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,587
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,492
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,598
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,613
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,582
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,596
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,612
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,632
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,601
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,580
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,568
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,314
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,569
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,311
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,671
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,674
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,669
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,312
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,313
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,672
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,315
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/056,886
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,877
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,889
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,893
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,630
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,878
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,662
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,872
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,882
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,637
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,903
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,888
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,879
```



[illegible]

```

; EARLIER APPLICATION NUMBER: 60/056,636
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,874
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,910
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,864
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,631
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,845
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,892
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/057,761
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/047,595
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,599
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,588
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,585
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,586
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,590
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,594
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,589
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,593
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,614
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,578
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,576
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/047,501
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,670
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/056,632
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,664
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,876
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,881
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,909
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,875
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,862
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,887
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,908
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/057,650
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/056,884
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/057,669
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/049,610
; EARLIER FILING DATE: 1997-06-13
; EARLIER APPLICATION NUMBER: 60/061,060

```

```

; EARLIER FILING DATE: 1997-10-02
Alignment Scores:
Pred. No.: 2,428-36 Length: 2664
Score: 659.50 Matches: 150
Percent Similarity: 58.75% Conservative: 48
Best Local Similarity: 44.51% Mismatches: 71
Query Match: 34.53% Indels: 69
DB: 3 Gaps: 9

US-09-544-776-2 (1-373) x US-09-149-476-255 (1-2664)
QY 42 GluGluAspGluAspGluLeuGluGluValLeuGluArgLysProAlaAla 61
DB 11 GAGCAGCGGAGTAAGGACTTGAGCGAG-----CCAGTTGCC 49
QY 62 GlyLeuSerAlaAlaProValProThrAlaProAlaAlaGlyAlaProLeuMetAspPhe 81
DB 50 GGATTATTCTATTTCCTCC----- 67
QY 82 GlyAsnAspPheValProAlaProAlaProArgGlyPheLeuProAlaAlaProPro----- 99
DB 68 -----TCCCTCTCTCCCGCCCGTATCTTTTACCTCTCTCCACCTCTCCACCTCGCTCG 118
QY 100 -----ValAlaProGluArgGlnProSerTrpAspProSerProValSerSerThrVal 117
DB 119 CGTASCATGGCGAGCGTGGCGGCA-----CTCAGTCCCAT 157
QY 118 ProAlaProSerPheLeuSerAlaAlaValSerProSerLysLeuProGluAspAsp 137
DB 158 CCATCTCTCT----- 166
QY 138 GluProProAlaArgProProProProProProAlaSerValSerProGlnAlaGluPro 157
DB 167 CQTCTCTCTCGGAGCGCGTCCCGCGCGGCGGCGGAGCCAG---GAGCCT 223
QY 158 ValTrpThrProAlaProAlaProAlaProAlaProSerThrProAlaAlaProLys 177
DB 224 -----GCCCGCGCC-----TGGGACGA 241
QY 178 ArgArgGlySerSerGlySerValValValAspLeuLeuTrpTrpArgAspLysLys 197
DB 242 AGAGCTGC-AGCTCCTCTCTGTCGGTGACGATCTGATTTCTGGAGAGATGAGAGAAG 300
QY 198 ThrGlyValValPheGlyAlaSerLeuPheLeuLeuSerLeuThrValPheSerIle 217
DB 301 ACTGGGTTTGTCTTTGGCACCACCGTATCATCTGCTTTCTCCCGCAGCTTTCAGTGTC 360
QY 218 ValSerValThrAlaTyrlleAlaLeuAlaLeuLeuSerValThrIleSerProArgIle 237
DB 361 ATCAGTGTGTTTCTTACCTCATCTGGCTCTTCTCTGTGCCATCAGCTTCAGGATC 420
QY 238 TyrLysGlyValIleGlnAlaIleGlnLysSerAspGluGlyHisProPheArgAlaTyr 257
DB 421 TACAAGTCGTTCATCCAAGCTGTACAGAAGTCAGAAGAAGGCCATCCATTCAAAGCCTAC 480
QY 258 LeuGluSerGluValAlaIleSerGluGluLeuValGlnLysTyrSerAsnSerAlaLeu 277
DB 481 CTGGACGTAGACATTACTCTGTCTCAGAAAGTTCCTATAATTACATGATGTCGCAATG 540
QY 278 GlyHisValAsnCysThrIleLysGluLeuArgArgLeuPheLeuValAspAspLeuVal 297
DB 541 GTGCACATCAACAGGGCCCTGAAACTCATTTATTCGTCTCTTTCTGGTAGAAGATCTGGTT 600
QY 298 AspSerLeuLysPheAlaValLeuMetTrpValPheThrTyrValGlyAlaLeuPheAsn 317
DB 601 GACTCCTTGAAGCTGGCTGTCTTCATGTGGCTGATGACCTATGTTGGTCTGTTTAAAC 660
QY 318 GlyLeuThrLeuLeuIleLeuAlaLeuIleSerLeuPheSerValProValIleTyrGlu 337
DB 661 GGAATCACCTTCTAATTTCTTCTGAACGTGCTCATTTTCAGTGTCCGAGTGTCTATGAG 720
QY 338 ArgHisGlnAlaGlnIleAspHisTyrLeuGlyLeuAlaAsnLysAsnValLysAspAla 357

```



```
Db 721 AATGACCAAGCCAGATTTGATCTATGTTGGCATCGCCGAGATCAGACCAAGTCAATT 780
Qy 358 MetAlaLysIleGlnAlaLysIleProGlyLeu---LysArgLysAlaGlu 373
Db 781 GTTGAAAGATCCAGCAAACTCCCTGGAAATGCCCAAAAAAAGGCAGAA 831

RESULT 9
US-09-949-016-2988
; Sequence 2988, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2988
; LENGTH: 2262
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-2988

Alignment Scores:
Pred. No.: 2,936-33 Length: 2262
Score: 613.50 Matches: 114
Percent Similarity: 80.42% Conservative: 38
Best Local Similarity: 60.32% Mismatches: 36
Query Match: 32.12% Indels: 1
DB: 4 Gaps: 1

US-09-544-776-2 (1-373) x US-09-949-016-2988 (1-2262)
Qy 186 ValValAspLeuLeuTyrTrpArgAspIleLysThrGlyValValPheGlyAlaSer 205
Db 10 GTCACGATCTGATTTCTGAGAGATGTGAGAGACTGGTGTGTTGTTGGCACCAG 69
Qy 206 LeuPheLeuLeuSerLeuThrValPheSerIleValSerValThrAlaTyrIleAla 225
Db 70 CTGATCATGCTGCTTTCCCTGGCAGCTTTTCAGTGTATCATGCTGTTCTTACCTCATC 129
Qy 226 LeuAlaLeuSerValThrIleSerProArgIleTyrLysGlyValIleGlnAlaIle 245
Db 130 CTGGCTCTTCTCTGTCACCATCAGCTTCAGATCTACAGTCCGTCATCCAGCTGTA 189
Qy 246 GlnLysSerAspGluGlyHisProPheArgAlaTyrIleuGluSerGluValAlaIleSer 265
Db 190 CAGAAGTCAGAGAAGGCCATCCATCAAGCCCTACCTGAGCAGTACAGATTAATCTGTCC 249
Qy 266 GluGluLeuValGlnLysTyrSerAsnSerAlaLeuGlyHisValAsnCysThrIleLys 285
Db 250 TCAGAAGCTTTCCATTAATACATGAATGTCGCCATCATCAATCAACAGGGCCCTGAAA 309
Qy 286 GluLeuArgArgLeuPheLeuValAspLeuValAspSerLeuLysPheAlaValLeu 305
Db 310 CTCAATTATTCCTCTCTTCTGTTAGAGATCTGTTGACTCTCTTGAAGCTGGCTGCTTC 369
Qy 306 MetTrpValPheThrTyrValGlyAlaLeuPheAsnGlyLeuThrIleuLeuAla 325
Db 370 ATGTGGCTGATGACCTATGTTGGTGTGTTTAAACGAATCACCTCTTAATCTTGTCT 429
Qy 326 LeuIleSerLeuPheSerValProValIleTyrGluArgHisGlnAlaGlnIleAspHis 345
Db 430 GAAGTGTCTATTTTCAGTGTCCGATTTGCTATGAGAAGTACAAAGACCCAGATTGATCAC 489
```

```
Qy 346 TyrLeuGlyLeuAlaAsnLysAsnValLysAspAlaMetAlaLysIleGlnAlaLysIle 365
Db 490 TATGTTGGCATCGCCGAGATCAGACCAAGTCAATTGTTGAAAGATCCAAGCAAACTC 549
Qy 366 ProGlyLeu---LysArgLysAlaGlu 373
Db 550 CCTGGAATCGCCAAAAAAGGCAGAA 576

RESULT 10
US-08-700-607-4
; Sequence 4, Application US/08700607
; Patent No. 5858708
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Au-Young, Janice
; APPLICANT: Goli, Surya K.
; APPLICANT: Hillman, Jennifer L.
; TITLE OF INVENTION: TWO NOVEL HUMAN NSP-LIKE PROTEINS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: U.S.
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/700,607
; FILING DATE: Filed Herewith
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0114 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1095 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; LIBRARY: THPINOBO1
; CLONE: 31870
US-08-700-607-4

Alignment Scores:
Pred. No.: 2,756-29 Length: 1095
Score: 550.00 Matches: 131
Percent Similarity: 58.78% Conservative: 43
Best Local Similarity: 44.26% Mismatches: 74
Query Match: 28.80% Indels: 49
DB: 2 Gaps: 7

US-09-544-776-2 (1-373) x US-08-700-607-4 (1-1095)
Qy 104 ArgGlnProSerTrpAspProSerProValSerSerThrVal---ProAlaProSerPhe 122
Db 21 CGAACCTCTAGTCGCGACCTCGGANTGAGTCAGTCAGTCGTCTCGGAGTC 80
Qy 123 LeuSerAlaAlaAlaValSerProSerLysLeuProGlu---AspAspGluProProAla 141
Db 81 AGCGGAGTAAGGAGCTTGAGCGACCGAGTTCGCGGATTAATTCATTTCCCTCTCT 140
Qy 142 ArgProPro----- 144
```



Db 3475 TGGCATCGCCGAGATCAGACCAAGTCAATTGTTGAAAG 3514  
RESULT 12  
US-09-149-476-102  
; Sequence 102, Application US/09149476  
; Patent No. 6420526  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: 186 Human Secreted proteins  
; FILE REFERENCE: P2002P1  
; CURRENT APPLICATION NUMBER: US/09/149,476  
; CURRENT FILING DATE: 1998-09-08  
; EARLIER APPLICATION NUMBER: PCT/US98/04493  
; EARLIER FILING DATE: 1998-03-06  
; EARLIER APPLICATION NUMBER: 60/040,162  
; EARLIER FILING DATE: 1997-03-07  
; EARLIER APPLICATION NUMBER: 60/040,333  
; EARLIER FILING DATE: 1997-03-07  
; EARLIER APPLICATION NUMBER: 60/038,621  
; EARLIER FILING DATE: 1997-03-07  
; EARLIER APPLICATION NUMBER: 60/040,626  
; EARLIER FILING DATE: 1997-03-07  
; EARLIER APPLICATION NUMBER: 60/040,334  
; EARLIER FILING DATE: 1997-03-07  
; EARLIER APPLICATION NUMBER: 60/040,336  
; EARLIER FILING DATE: 1997-03-07  
; EARLIER APPLICATION NUMBER: 60/040,163  
; EARLIER FILING DATE: 1997-03-07  
; EARLIER APPLICATION NUMBER: 60/047,600  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,633  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,583  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,617  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,618  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,503  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,592  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,581  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,584  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,500  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,587  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,492  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,598  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,613  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,582  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,596  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,612  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,632  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,601  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/043,580

; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,568  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,314  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,569  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,311  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,671  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,674  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,669  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,312  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,313  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,672  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,315  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/048,974  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/056,886  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,877  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,889  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,893  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,630  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,878  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,662  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,872  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,882  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,637  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,903  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,888  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,879  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,880  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,894  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,911  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,636  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,874  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,910  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,864  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,631  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,845  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,892  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/057,761  
; EARLIER FILING DATE: 1997-08-22



```

; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1419
; LENGTH: 2181
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-1419

Alignment Scores:
Pred. No.:      2,44e-27      Length:      2181
Score:          527.50      Matches:      149
Percent Similarity: 43.74%      Conservative: 57
Best Local Similarity: 31.63%      Mismatches:  140
Query Match:      27.62%      Indels:      125
DB:              4          Gaps:         15

US-09-544-776-2 (1-373) x US-09-949-016-1419 (1-2181)

Qy 17 ProArgProGlnProAlaPheLeuTyxGlnPheValArgGluProGluAspGluGlu 36
Db 385 CCCC GCCC CCCC CAGGCGC-----CGCTCAGTCTCGGAACCAACGAGAC----- 423
Qy 37 GluGluGluGluGluGluAepGluAepGluAepLeuGluGluLeuGluValLeuGlu 56
Db 424 -----CAGCACCCCTCAGCCAGCCTTGGCCACAGACTTGGAGAGCATCCCCAGCCTGAGC 477
Qy 57 ArgLysProAlaAlaGlyLeuSerAlaAlaProValProThrAlaProAlaAlaGlyAla 76
Db 478 CAATCCCGGAGCCTCGGACGAGCGGGGTGATCCT---GACACCGCGCTTCATCCGAGCGC 534
Qy 77 ProLeuMetAepPheGly-----AenAepPheValProProAla 89
Db 535 CCTCTGGAAGACCTGAGGCTTCGGTTGGACCATCTGGGCTGGGTGGCCCGGGGAACGGGA 594
Qy 83 -----TCCGGGGAGGACTCTTCCACCAGCAGCTCCACCCCGCTGGAAGACGAAGAACCCCAAGAA 654
Db 595 TCCGGGGAGGACTCTTCCACCAGCAGCTCCACCCCGCTGGAAGACGAAGAACCCCAAGAA 654
Qy 90 ProArgGlyPhe-----LeuProAlaAla 97
Db 655 CCCAACAGATTGGAGACAGGAGAGAGCTGGGGGAAGACTGGACCTACGACTCCGACTTGCT 714
Qy 98 ProProValAlaProGlu---ArgGlnProSerTrpAepPro-----SerProVal 113
Db 715 CAGCCCTCATGCCCGAGGCTTGACTCCCGAGCTCAGTCGGGCTCTGGGACACCCCGAG 774
Qy 114 SerSerThrValProAlaProSerPheLeuSerAlaAlaValSerProSerLys--- 132
Db 775 GCCGGTACT---CGGTCCCATCCGATCGGAGATTGGAATCTGGGCCCGGAAGAGCCA 831
Qy 133 LeuProGluAepAepGluProProAlaAargProProProProProProAla----- 149
Db 832 TTGCTGGAAGAGGAAGAAAGACAGTGGGGGCCCTTGGAGCGAGAGCCAGTAAGGGGACAG 891
Qy 150 -----SerValSerProGln----- 154
Db 892 TGCCTCATAGCAGGACCAATTAGAATTACGGTGGAGGCCCGCCTTCTAGGAACAGCT 951
Qy 154 ----- 154
Db 952 ATGGAATGGTTAAAGACATCATTCGTTTGGCTGTTTAAAGACGGTTTCCAATTTGGAA 1011
Qy 155 ---AlaGluProValTrpThr-----ProProAlaProAla 165
Db 1012 TTGTGCCCACTCTGTGGACAGCAATTGGCTGGGTGCCAAAGGGGGCCCAACCCCTTACT 1071

```

DB:	4	Gaps:	0
US-09-544-776-2 (1-373) x US-09-270-767-13561 (1-2014)			
Qy	179	AtqGlySerSerGlySerValValAspLeuLeuTyrTrpArgAspIleLeuLysLysThr	198
Db	134	CGMAAAATCTCGAACAGATGTTGGATCCCTTATCTACTGGCGCGATGTGAAGAATCC	193
Qy	199	GlyValValPheGlyAlaSerLeuPheLeuLeuLeuSerLeuThrValPheSerIleVal	218
Db	194	GGCAVTTGTTCTTCGGCGCTGGCCTGATCACATGTCGGCGCATCTCCAGCTTCTCGGTGATC	253
Qy	219	SerValThrAlaTyrIleAlaLeuAlaLeuSerValThrIleSerProArgIleTyr	238
Db	254	AGCGTGTTCGCCCTACTTTCGCTCTCAACCTCTTCGGCACCGTCGCTTCAGAATCTAC	313
Qy	239	LysGlyValIleGlnAlaIleGlnLysSerAspGluGlyHisProPheArgAlaTyrLeu	258
Db	314	AAATCTGTGACACAGCGCTGCMAAAGACAACAGAGGTCAACCCCTTAAGGATTACCTG	373
Qy	259	GluSerGluValAlaIleSerGluGluLeuValGlnLysTyrSerAsnSerAlaLeuGly	278
Db	374	GAGCTGGATCTGACGCTGTGCGCACGAAAGGTACAGAACATTCGCGCGTGGCTGTGGCA	433
Qy	279	HisValAsnCysThrIleLysGluLeuArgGluPheLeuLeuAlaAspLeuValAsp	298
Db	434	CATATCAATGGCTTCATCTCCAGCTGAGGCGTCTGTTCTTGTGTAGGATATCATCGAT	493
Qy	299	SerLeuLysPheAlaValLeuMetTrpValPheThrTyrValGlyAlaLeuPheAsnGly	318
Db	494	TCGATCAAGTTCGGCGTCATCTGTGGGTCTTCACCTACGTGGGTGCTCGTTCAATGGC	553
Qy	319	LeuThrLeuLeuIleAlaLeuLleSerLeuPheSerValProValIleTyrGluArg	338
Db	554	ATGACTCTGGTCATCTTGGCCTTTGTCTCGCTGTTTACCTTGTGCCAAGGTCTACGAGAAC	613
Qy	339	HisGlnAlaGlnIleAspHisTyrLeuGlyLeuAlaAsnLysAsnValLysAspAlaMet	358
Db	614	AACAGCAATCGATCGACACTCACTTGGATCTGGTGGCGAGCAAAATGACAGAAATCACC	673
Qy	359	AlaLysIleGlnAlaLysIleProGlyLeuLysArgLysAlaGlu	373
Db	674	GACAAGATCCGAGTGGCATCCCATTCGTCACCAAGAAGCCCGAG	718

RESULT 15

```

US-09-513-999C-2227
; Sequence 2227, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 2227
; LENGTH: 441
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 236..439
; US-09-513-999C-2227

```

Alignment Scores:	
Pred. No.:	8.77e-16
Score:	347.00
	Length: 441
	Matches: 68

```

Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match: 18.17%              Indels: 0
DB: 4                             Gaps: 0

US-09-544-776-2 (1-373) x US-09-513-999C-2227 (1-441)

Qy 1 MetGluAspLeuAspGlnSerProLeuValSerSerAspSerProProArgProGln 20
    |||
Db 236 ATGGAGACCTGGACAGTCCTCTGTCCTCGGACAGCCACCCCGCCGCGAG 235
    |||

Qy 21 ProAlaPheLysTyrGlnPheValArgGluProGluAspGluGluGluGluGlu 40
    |||
Db 296 CCGGGTTCAAGTACCACTTCGTGGAGAGCCCGAGCGAGGAGAGAGAGGAGGAG 355
    |||

Qy 41 GluGluGluAspGluAspGluAspGluLeuGluLeuGluValLeuGluArgLysProAla 60
    |||
Db 356 GAAGAGGAGGACGAGGACGAAGACCTGGAGGAGCTGGAGGTGCTGGAGAGAGCCCGCC 415
    |||

Qy 61 AlaGlyLeuSerAlaAlaProVal 68
    |||
Db 416 GCCGGCTGTCCGGCGCCCACTG 439
    |||

Search completed: August 3, 2005, 23:15:46
Job time : 224 secs

```

Search completed: August 3, 2005, 23:15:46  
Job time : 224 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 3, 2005, 17:14:22 ; Search time 4328 Seconds  
(without alignments)  
4176.016 Million cell updates/sec

Title: US-09-544-776-2  
Perfect score: 1910  
Sequence: 1 MEDLSDPLVSSSDSPRPQ.....VKDAMAKIOAKIPGLKRAE 373

Scoring table: BLOSUM62 Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 1839042

Minimum DB seq length: 0  
Maximum DB seq length: 50

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlp  
-Q=/cgn2\_1/USPTO.spool\_p/US09544776/runat\_03082005\_181414\_27618/app.query.fasta\_1.519  
-DB=GenEmbl -QFWT=fastap -SUFFIX=p2nslm50.rge -MINMATCH=0.1 -LOOPL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFWT=ptp -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=50  
-USER=US09544776 @CGN 1.1 5600 @runat\_03082005\_181414\_27618 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.\*  
1: gb.ba.\*  
2: gb.htg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.sts.\*  
12: gb.sy.\*  
13: gb.un.\*  
14: gb.vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	68	3.6	48	6 AR199544	Sequence
2	68	3.6	48	6 AR374696	Sequence
3	68	3.6	48	6 AR409331	Sequence
4	68	3.6	48	6 AX239593	Sequence

5	68	3.6	48	6	BD056458	Novel low
6	66	3.5	48	6	AR178317	Sequence
7	66	3.5	48	6	AX323399	Sequence
8	66	3.5	50	6	AS1711	Sequence 17
9	66	3.5	50	6	AR167590	Sequence
10	66	3.5	50	6	AR178300	Sequence
11	66	3.5	50	6	AX323382	Sequence
12	66	3.5	50	6	AX686852	Sequence
13	63	3.3	50	6	AR356035	Sequence
14	63	3.3	50	6	AR537591	Sequence
15	60	3.1	50	6	AR232180	Sequence
16	60	3.1	50	6	AX539582	Sequence
17	60	3.1	50	6	AX539583	Sequence
18	59	3.1	45	6	AR199538	Sequence
19	59	3.1	45	6	AR374690	Sequence
20	59	3.1	45	6	AR409325	Sequence
21	59	3.1	45	6	AX239587	Sequence
22	59	3.1	45	6	BD056452	Novel low
23	58	3.0	34	6	BD169737	Sequence
24	58	3.0	38	6	BD169738	Sequence 1
25	58	3.0	40	6	BD169738	C-terminu
26	57	3.0	39	6	BD169738	C-terminu
27	57	3.0	42	6	BD247090	Improved
28	57	3.0	47	6	AX539586	Sequence
29	57	3.0	47	6	AX539587	Sequence
30	57	3.0	48	6	AR178317	Sequence
31	57	3.0	48	6	AR306664	Sequence
32	57	3.0	48	6	AR340099	Sequence
33	57	3.0	48	6	AR412168	Sequence
34	57	3.0	48	6	AX323399	Sequence
35	57	3.0	50	6	AS1711	Sequence 17
36	57	3.0	50	6	AR167590	Sequence
37	57	3.0	50	6	AR178300	Sequence
38	57	3.0	50	6	AX323382	Sequence
39	57	3.0	50	6	AX686852	Sequence
40	56	2.9	42	6	AX512498	Sequence
41	56	2.9	42	6	AX573414	Sequence
42	56	2.9	42	6	AX746300	Sequence
43	56	2.9	48	6	AR306674	Sequence
44	56	2.9	48	6	AR340109	Sequence
45	56	2.9	48	6	AR412178	Sequence

ALIGNMENTS

RESULT 1  
AR199544  
LOCUS AR199544 48 bp DNA linear PAT 20-APR-2002  
DEFINITION Sequence 36 from patent US 6355451.  
ACCESSION AR199544  
VERSION AR199544.1 GI:20249618  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 48)  
AUTHORS Lees,A.M., Lees,R.S., Law,S.W. and Arjona,A.A.  
TITLE Low density lipoprotein binding proteins and their use in  
diagnosing and treating atherosclerosis  
JOURNAL Patent: US 6355451-A 36 12-MAR-2002;  
FEATURES  
source Location/Qualifiers  
1..48  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN

Alignment Scores:  
Pred. No.: 3.39e+04 Length: 48  
Score: 68.00 Matches: 11  
Percent Similarity: 100.00% Conservative: 5  
Best Local Similarity: 68.75% Mismatches: 0  
Query Match: 3.56% Indels: 0  
DB: 6 Gaps: 0

US-09-544-776-2 (1-373) x AR409331 (1-48)

**FEATURES**  
**SOURCE**





[illegible]

```

source      1..50
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Alignment Scores:
Pred. No.:      5.59e+04      Length:      50
Score:          63.00         Matches:    12
Percent Similarity: 46.15%    Conservative: 0
Best Local Similarity: 46.15% Mismatches:    2
Query Match:     3.30%       Indels:      12
DB:              6           Gaps:        1

US-09-544-776-2 (1-373) x AR356035 (1-50)

Qy      139  ProProAlaArgProProProProProAlaSerValSerProGlnAlaGluProVal 158
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      5   CCTCCCCCTCCCNCCNCCCCCCCCCCCCCCCC----- 34

Qy      159  TrpThrProProAlaPro 164
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      35  -----CCCCCGGCCCC 46

RESULT 14
AR537591
LOCUS      AR537591
DEFINITION Sequence 2153 from patent US 6737248.
ACCESSION AR537591
VERSION   AR537591.1 GI:53928808
KEYWORDS

```

RECEIVED  
VERSION AR537591.1 GI:53928808  
KEYWORDS

REFERENCE  
1. (bases 1 to 50)  
Kunsch, C.A., Choi, G.A., Barash, S.C., Dillon, P.J., Fannon, M.R. and  
rosen, C.A.  
TITLE  
Scaphylococcus aureus polynucleotides and sequences  
JOURNAL  
Patent: US 6737248-A 2153 18-MAY-2004;  
FEATURES  
Location/Qualifiers  
1. .50  
source  
/organism="unknown"  
/mol type="genomic DNA"

[illegible]

Qy	159	TpThrProAlaPro	164
Db	35	-----CCCCCGGCCCCC	46
RESULT	15		
AR232180/c			
LOCUS	AR232180	50 bp	DNA
DEFINITION	Sequence 14 from patent US 6455392.		
ACCESSION	AR232180		
VERSION	AR232180.1	GI:27274071	
KEYWORDS	.		
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	Unclassified.		
	1 (bases 1 to 50)		

**THIS PAGE BLANK (U&PTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 3, 2005, 17:14:22 ; Search time 560 Seconds  
(without alignments)  
3942.971 Million cell updates/sec

Title: US-09-544-776-2

Perfect score: 1910

Sequence: 1 MEDLDOSPLVSSSDSPRPQ.....VKDAMAKIQAKIPGLKRAE 373

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 4167226

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DRV=xlp  
-Q=/cd2\_1/USPTO\_spool\_p/US09544776/runat\_03082005\_181413\_27611/app.query.fasta\_1.519  
-DB=N\_Geneseq\_16Dec04 -QFMT=fastap -SUFFIX=p2nslm50.rng -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0  
-MAXLEN=50 -USER=US09544776 @CGN 1 1 708 @runat\_03082005\_181413\_27611 -NCFU=6  
-ICPU=3 -NO\_WMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N\_Geneseq\_16Dec04:\*

- 1: Geneseqn1980s:\*
- 2: Geneseqn1990s:\*
- 3: Geneseqn2000s:\*
- 4: Geneseqn2001as:\*
- 5: Geneseqn2001bs:\*
- 6: Geneseqn2002as:\*
- 7: Geneseqn2002bs:\*
- 8: Geneseqn2003as:\*
- 9: Geneseqn2003bs:\*
- 10: Geneseqn2003cs:\*
- 11: Geneseqn2003ds:\*
- 12: Geneseqn2004as:\*
- 13: Geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	68	3.6	48	5	Aah26508
2	66	3.5	48	6	Aas21106 (GGA)16 D
3	66	3.5	50	2	Aat32776 Homopurin
4	66	3.4	42	10	Aas19342 Plasmid X
5	64	3.4	42	10	Abz22128 Polyanion

64	3.4	42	10	Abz222092
64	3.4	45	4	Aai64451 SSR motif
64	3.4	50	10	Adc56858 Micro gol
64	3.4	50	10	Adc56859 Micro gol
63	3.3	50	2	Aav76464 Staphyloc
60	3.1	50	6	Abse67031 Human MRP
60	3.1	50	6	Abse67032 Human MRP
60	3.1	50	8	Abx13592 Human ser
60	3.1	50	10	Adh59088 Human KSE
60	3.1	50	12	Adh59088 Human KSE
59	3.1	42	4	Aas13776 Simple se
59	3.1	42	4	Aas13779 Simple se
59	3.1	45	5	Aah26502 Low densi
58	3.0	34	6	Abk95216 PCR prime
58	3.0	38	2	Aat06967 Primer fo
58	3.0	40	6	Abk95217 PCR prime
57	3.0	39	12	Adk61703 Base cont
57	3.0	47	6	Abse67036 Human MRP
57	3.0	47	6	Abse67035 Human MRP
57	3.0	48	2	Aax27671 [GPP]in 1
57	3.0	48	6	Aas21106 (GGA)16 D
57	3.0	48	6	Abk95217 PCR prime
57	3.0	48	6	Abk95217 PCR prime
57	3.0	50	2	Aat32776 Homopurin
57	3.0	50	6	Aas19342 Plasmid X
56	2.9	42	4	Aah48668 Anti-trop
56	2.9	42	5	Aah21345 CSC-F ove
56	2.9	42	8	Abz21455 Chicken i
56	2.9	43	2	AAQ81075 supf gene
56	2.9	43	2	AAQ81075 supf gene
56	2.9	43	2	AAQ81075 supf gene
56	2.9	48	2	Aax27675 Potato Le
56	2.9	48	6	Aas16547 Ligatoin
56	2.9	48	6	Abk95217 PCR prime
56	2.9	48	6	Abk95217 PCR prime
55	2.9	50	10	ADD31857 Synthetic
55	2.9	45	2	AAV5838 Sequence
55	2.9	37	2	AAQ05014 Sequence
55	2.9	37	2	AAQ36334 HSPOL2pa
55	2.9	37	2	AAQ36335 HSPOL2an
55	2.9	38	6	Abse67040 Human MRP
55	2.9	38	6	Abse67039 Human MRP
55	2.9	40	3	Aaz96053 Polynucle

ALIGNMENTS

RESULT 1  
AAH26508  
ID AAH26508 standard; DNA; 48 BP.

XX AAH26508;

DT 12-NOV-2001 (first entry)

XX Low density lipoprotein binding protein (LBP) polynucleotide.

XX Low density lipoprotein binding protein 1; LBP-1; LDL; atherosclerosis;  
XX antiarteriosclerotic; therapy; diagnosis; vaccine; ss.

XX Mammalia.

XX WO200164874-A2.

XX 07-SEP-2001.

XX 28-FEB-2001; 2001WO-US006356.

XX 02-MAR-2000; 2000US-00517849.

XX 14-JUL-2000; 2000US-00616289.

XX (BOST-) BOSTON HEART FOUND INC.

XX Lees AM, Lees RS, Law SW, Arjona AA;

XX WPI; 2001-565505/63.

















```

XX
SQ Sequence 50 BP; 2 A; 17 C; 30 G; 1 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 9.67e+04 Length: 50
Score: 60.00 Matches: 12
Percent Similarity: 46.15% Conservative: 0
Best Local Similarity: 46.15% Mismatches: 2
Query Match: 3.14% Indels: 12
DB: 12 Gaps: 1

```

US-09-544-776-2 (1-373) x ADH59088 (1-50)

Qy 143 ProProProProProProAlaSerValSerProGlnAlaGluProValTrpThrProPro 162  
 Db 42 CCGCGCGCGCGCGCGCG-----CCGCGG 19  
 Qy 163 AlaProAlaProAlaAla 168  
 Db 18 AGGCCCGCGCGCCCGCGCC 1

Search completed: August 4, 2005, 01:11:41  
Job time : 563 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 3, 2005, 17:14:22 ; Search time 3086 Seconds  
(without alignments)  
4600.769 Million cell updates/sec

Title: US-09-544-776-2

Perfect score: 1910

Sequence: 1 MEDLOSPLVSSSDSPPRPQ.....VKDAMAKIQAKIPGLKRRAE 373

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Xgapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 159776

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp  
-Q/cgn2\_1/USPTO\_spool\_p/US09544776/runat\_03082005\_181414\_27631/app\_query.fasta\_1.519  
-DB=EST -Qfmt=fastap -SUFFIX=p2nslm50.rst -MINMATCH=0.1 -LOOPCL=0 -LOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=50  
-USER=US09544776 @CGN 1.1 5180 @runat\_03082005\_181414\_27631 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST: \*  
1: gb\_est1: \*  
2: gb\_est2: \*  
3: gb\_hic: \*  
4: gb\_est3: \*  
5: gb\_est4: \*  
6: gb\_est5: \*  
7: gb\_est6: \*  
8: gb\_gest1: \*  
9: gb\_gest2: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	70	3.7	49	1 AA145482	AA145482 mr62a06.r
2	70	3.7	50	9 BX976589	BX976589 Forward s
3	68	3.6	50	9 CG731182	CG731182 1119136A0
4	66	3.5	49	8 AZ840678	AZ840678 2M0138016
5	66	3.5	50	9 CR041958	CR041958 Forward s
6	65	3.4	31	8 AZ861612	AZ861612 2M0168P16
7	65	3.4	48	9 AZ752522	AZ752522 Arabidops
8	65	3.4	49	9 TA33D12P	TA33D12P T. brucei
9	65	3.4	50	9 BX002655	BX002655 Arabidops

C	10	65	3.4	50	9	CR154670	CR154670 Reverse s
	11	64	3.4	48	8	AZ646447	AZ646447 1M0512A06
	12	64	3.4	49	8	AZ506149	AZ506149 1M0347J05
	13	64	3.4	49	8	AZ656875	AZ656875 1M0532M11
	14	64	3.4	49	8	AZ764533	AZ764533 1M0560Q17
	15	64	3.4	49	8	AZ773388	AZ773388 1M0584E23
	16	64	3.4	49	8	AZ820100	AZ820100 2M0092A12
	17	64	3.4	49	9	TA168G09Q	TA168G09Q
	18	64	3.4	49	4	BI491464	BI491464 df09f10.w
	19	64	3.4	50	7	CV304274	CV304274 t594902.b
	20	64	3.4	50	8	AZ358097	AZ358097 1M0100N08
	21	64	3.4	50	8	AZ456166	AZ456166 1M0258B20
	22	64	3.4	50	8	AZ787456	AZ787456 2M0033O06
	23	64	3.4	50	8	AZ816605	AZ816605 2M0085K24
	24	64	3.4	50	8	AZ861629	AZ861629 2M0168A19
	25	64	3.4	50	8	AZ967786	AZ967786 2M0238F20
	26	64	3.4	50	9	BX987794	BX987794 Reverse s
	27	64	3.4	50	9	CR03825	CR03825 Forward s
	28	64	3.4	50	9	CR025955	CR025955 Reverse s
	29	64	3.4	50	9	CR033162	CR033162 Forward s
	30	64	3.4	50	9	CR048038	CR048038 Reverse s
	31	64	3.4	50	9	CR085382	CR085382 Reverse s
	32	64	3.4	50	9	CR122615	CR122615 Forward s
	33	64	3.4	50	9	CR166872	CR166872 Forward s
	34	63	3.3	49	1	AI931428	AI931428 ul68C04.Y
	35	63	3.3	49	8	AZ335579	AZ335579 1M0065A01
	36	62	3.2	39	8	AZ826077	AZ826077 2M0101A07
	37	62	3.2	42	9	TA63A10P	TA63A10P
	38	62	3.2	44	8	AZ342202	AZ342202 1M0075G11
	39	62	3.2	46	8	AZ981881	AZ981881 2M0262P23
	40	62	3.2	50	9	CR155014	CR155014 Forward s
	41	61	3.2	48	8	AZ652813	AZ652813 1M0526N08
	42	60.5	3.2	50	9	CR117295	CR117295 Reverse s
	43	60	3.1	33	8	AZ317100	AZ317100 1M0035E01
	44	60	3.1	35	8	AZ369295	AZ369295 1M0119M13
	45	60	3.1	37	8	AZ369295	AZ369295 1M0119M13

ALIGNMENTS

RESULT 1  
AA145482  
LOCUS  
DEFINITION  
AA145482 49 bp mRNA linear EST 11-FEB-1997  
mr62a06.r1 Stratagene mouse testis (#937308) Mus musculus CDNA  
clone IMAGE:602002 5', similar to TR:G307311 G307311 HISTONE H4. ;,  
mRNA sequence.

ACCESSION  
AA145482 GI:1714857

VERSION  
AA145482.1

KEYWORDS  
EST.

SOURCE  
Mus musculus (house mouse)

ORGANISM  
Mus musculus

REFERENCE  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

AUTHORS  
Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisler, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.

TITLE  
The WashU-HMI Mouse EST Project

JOURNAL  
Unpublished (1996)

COMMENT  
Contact: Marra M/Mouse EST Project  
WashU-HMI Mouse EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: mouseest@watson.wustl.edu  
This clone is available royalty-free through LLNL ; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
MGI:367434  
Possible reversed clone: similarity on wrong strand  
Seq primer: -28ml3 rev1 ET from Amersham







```

84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0168 row: p column: 16
Seq primer: CGTTGTAAACGACGCGCCACT
Class: plasmid ends
High quality sequence stop: 31.
    Location/Qualifiers
1. . 31
    /organism="Mus musculus"
    /mol_type="genomic DNA"
    /strain="C57BL/6J"
    /db_xref="taxon:10090"
    /clone="TUGC2M0168P16"
    /sex="Male"
    /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
    /clone_lib="Mouse 10kb plasmid UUGCIM library"
    /note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydronamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptored DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PWD42 [G14732114|gb|AF129072.1], a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adaptored mouse DNA was annealed to
adaptored vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

```

## ORIGIN

Alignment Scores:	
Pred. No.:	2.54e+04
Score:	65.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	3.40%
DB:	8
Length:	31
Matches:	10
Conservative:	0
Mismatches:	0
Indels:	0
Gaps:	0

US-09-544-776-2 (1-373) x AZ861612 (1-31)

Qy	139	ProProAlaArgProProProProPro	148
pb	2	CCCCCGGCGGCCCCCCCCCCCCCCCC	31

## RESULT 7

AL752522/c  
LOCUS  
DEFINITION  
AL752522  
48 bp DNA linear  
GSS 31-MAR-2004  
Arabidopsis thaliana T-DNA flanking sequence GK-014A05-012792,  
genomic survey sequence.

ACCESSION AL752522.1  
VERSION AL752522.1 GI:21485020

**KEYWORDS:** GSS.

**SOURCE** *Arabidopsis thaliana* (thale cress)

ORGANISM *Arabidopsis thaliana*

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.

1. REFERENCE  
AUTHORS Li, Y., Rosso, M.G., Strizhov, N., Viehoever, P. and Weisshaar, B.  
TITLE GABI-Kat SampleSearch: a flanking sequence tag (FST) database for  
the identification of T-DNA insertion mutants in Arabidopsis  
thaliana

JOURNAL Bioinformatics 19 (11), 1441-1442 (2003)

84112, USA	MEDLINE
Tel: 801 585 5606	PUBMED
Fax: 801 585 7177	REFERENCE
Email: ddunn@genetics.utah.edu	AUTHORS
Insert Length: 10000	Std Error: 0.00
Plate: 0168	row: P
	column: 16
Seq primer: CGTGTAAACGACGCCAGT	TITLE
Class: plasmid ends	JOURNAL
High quality sequence stop: 31.	MEDLINE
Location/Qualifiers	PUBMED
1. .31	REFERENCE
/organism="Mus musculus"	AUTHORS
source	

TITLE	JOURNAL	PUBMED	REFERENCE	AUTHORS	TITLE	JOURNAL
strains "C57BL/6J" and "129/Ola" were used						
lab hosts="E. Coli strain XL10-Gold, Tl-resistant, F-"						
clone lib="Mouse 10kb plasmid UUGC1M library"						
/note=vector: pWD42nv; Purified genomic DNA from M.						
musculus C57BL/6J (male) was obtained from the Jackson						
sex="Male"						
clone="UUGC2M0168P16"						

laboratory mouse DNA resources  
(<http://www.jax.org/resources/documents/dnares/>). The DNA

T-DNA. Details on the protocols used for generation of the sequences are described in References 1-3. Re-examination of the source from which this sequence has been produced indicates that the sequence is of low reliability. Therefore, no information on a potential insertion site is deduced. The sequences are generated at the MPI for Plant Breeding Research in the context of the GABI-Kat project. GABI-Kat is part of the German Plant Genomics program designated 'GABI'. Information on line availability can be found at: <http://www.mpiz-koeln.mpg.de/GABI-Kat/>.

inducible derivative of plasmid R1. The vector was ligated

**source**

/organism="Arabidopsis thaliana"

/mol\_type="genomic DNA"

```
/strain="Columbia 0"
```

/db\_xref="taxon:3702"

```
/clone="GK-014A05-012792"
```

```
/clone_lib="Arab
```

/note="PCR was performed on DNA from *Arabidopsis thaliana* plants (r1) which were transformed with the T-DNA from vector pC106 (Genbank accession number: AJ537513). The lines contain one or more T-DNA insertions. The DNA fragment(s) resulting from the PCR were directly sequenced to determine the genomic sequence flanking the insertion. T-DNA derived sequences were removed."

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

## ORIGIN

Alignment Scores:		
Pred. No.:	4.33e+04	48
Score:	65.00	12
Percent Similarity:	42.86%	0
Best Local Similarity:	42.86%	Conservative: 0
Query Match:	3.40%	Mismatches: 4
DB:	9	Indels: 12
		Gaps: 1
		Length: 48
		Matches: 12

US-09-544-776-2 (1-373) x AL752522 (1-48)

Qy	143	ProProProProProAlaSerValSerProGlnAlaGluProValTyrThrProPro	162
Db	48	CSCCCNCCCCCNCNCCC-----CCCCC	25

0v 163 A\aproA\aproA\apropro 170

24 CCCCCCNCCCCCCCCCCCCCCCCC 1

## RESULT 8

TA33D12P

LOCUS TA33D12P 49 bp DNA linear GSS 13-DEC-2000  
DEFINITION T. brucei sheared genomic DNA clone 33d12, forward sequence, genomic survey sequence.

```

ACCESSION AL453367
VERSION AL453367.1 GI:11854693
KEYWORDS GSS.
SOURCE Trypanosoma brucei
ORGANISM Trypanosoma brucei
          Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
          Trypanosoma.
REFERENCE 1 (bases 1 to 49)
AUTHORS Hall,N., Bowman,S., Lennard,N.J., Doggett,J., Atkin,R.,
          Chillingworth,C., Ormond,D., Harris,B., El-Sayed,N., Hou,L.,
          Melville,S.B., Rajandream,M.A. and Barrell,B.G.
          Direct Submission
          Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing
          project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton,
          Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and
          nh@sanger.ac.uk
TITLE Constructed at the Institute for Genomic Research (TIGR),
JOURNAL Rockville, MD. Genomic DNA isolated from a cloned population of
MEDLINE Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared
PUBMED to give a tight size distribution (
REFERENCE 4 kb). The v + i method used for the library construction is
AUTHORS described in detail in Smith, H. and Venter, J.C. (Making small
          insert libraries for whole genome shotgun sequencing projects. In
          Genome Sequencing: A Practical Approach, eds. M. Vaudin and B.
          Barrell, Oxford University Press, 1999).
          Email: nelsayed@tigr.org
COMMENT Details of T. brucei sequencing at the Sanger Centre are available
          at http://www.sanger.ac.uk/Projects/T_brucei/.
FEATURES             source
                     1..49
                     /organism="Trypanosoma brucei"
                     /mol_type="genomic DNA"
                     /strain="TREU927"
                     /db_xref="taxon:5691"
                     /clone="33d12"
ORIGIN
Alignment Scores:
Pred. No.:      4.44e+04      Length:      49
Score:          65.00      Matches:      12
Percent Similarity: 42.86%      Conservative: 0
Best Local Similarity: 42.86%      Mismatches: 4
Query Match:     3.40%      Indels:    12
DB:              9      Gaps:        1
US-09-544-776-2 (1-373) x TA33D12P (1-49)
QY 143 ProProProProProProAlaAlaProPro 170
Db 1 CCCCCCCCCCCCCCCCCC-----CCCCC 24
QY 163 AlaProAlaProAlaAlaProPro 170
Db 25 NNCNCNCCCCCCCCCCCCCCCCCCCC 48
RESULT 9
BX002655/c 50 bp DNA linear GSS 01-APR-2004
LOCUS Arabidopsis thaliana T-DNA flanking sequence GK-366A11-017045,
DEFINITION genomic survey sequence.
ACCESSION BX002655
VERSION BX002655.1 GI:26187615
KEYWORDS GSS.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
          rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE 1
AUTHORS Li,Y., Rosso,M.G., Strizhov,N., Viehoveer,P. and Weissshaar,B.
TITLE GABI-kat SimpleSearch: a flanking sequence tag (FST) database for
          the identification of T-DNA insertion mutants in Arabidopsis
          thaliana

```

```

JOURNAL Bioinformatics 19 (11), 1441-1442 (2003)
MEDLINE 22755829
PUBMED 12874060
REFERENCE 2
AUTHORS Rosso,M.G., Li,Y., Strizhov,N., Reiss,B., Dekker,K. and
          Weissshaar,B.
TITLE An Arabidopsis thaliana T-DNA mutagenized population (GABI-Kat) for
          flanking sequence tag-based reverse genetics
JOURNAL Plant Mol. Biol. 53 (1-2), 247-259 (2003)
MEDLINE 23117147
PUBMED 14756321
REFERENCE 3
AUTHORS Strizhov,N., Li,Y., Rosso,M.G., Viehoveer,P., Dekker,K.A. and
          Weissshaar,B.
TITLE High-throughput generation of sequence indexes from T-DNA
          mutagenized Arabidopsis thaliana lines
JOURNAL Biotechniques 35 (6), 1164-1168 (2003)
PUBMED 14682050
REFERENCE 4 (bases 1 to 50)
AUTHORS Strizhov,N., Li,Y., Rosso,M.G. and Weissshaar,B.
TITLE Direct Submission
JOURNAL Submitted (31-MAR-2004) Weissshaar B., Max-Planck-Institut fuer
          Zuechtungsforschung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany
          This sequence has been recovered from the left border of the T-DNA.
          Details on the protocols used for generation of the sequence are
          described in References 1-3. Re-examination of the source from
          which this sequence has been produced indicates that the sequence
          is of low reliability. Therefore, no information on a potential
          insertion site is deduced. The sequences are generated at the MPI
          for Plant Breeding Research in the context of the GABI-Kat project.
          GABI-Kat is part of the German Plant Genomics program designated
          'GABI'. Information on line availability can be found at:
          http://www.mpiz-koeln.mpg.de/GABI-Kat/.
FEATURES             source
                     1..50
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /strain="Columbia 0"
                     /db_xref="taxon:3702"
                     /clone="GK-366A11-017045"
                     /clone_lib="Arabidopsis thaliana T-DNA insertion lines"
                     /ecotype="Col-0"
                     /notes="PCR was performed on DNA from Arabidopsis thaliana
                     plants (Ti) which were transformed with the T-DNA from
                     vector pAC161 (GenBank accession number: AJ537514). The
                     lines contain one or more T-DNA insertions. The DNA
                     fragment(s) resulting from the PCR were directly sequenced
                     to determine the genomic sequence flanking the insertion.
                     T-DNA derived sequences were removed."
ORIGIN
Alignment Scores:
Pred. No.:      4.55e+04      Length:      50
Score:          65.00      Matches:      12
Percent Similarity: 42.86%      Conservative: 0
Best Local Similarity: 42.86%      Mismatches: 4
Query Match:     3.40%      Indels:    12
DB:              9      Gaps:        1
US-09-544-776-2 (1-373) x BX002655 (1-50)
QY 143 ProProProProProProAlaAlaProPro 162
Db 48 CCCCCCCCCCCCCCCCCC-----CCCCC 25
QY 163 AlaProAlaProAlaAlaProPro 170
Db 24 CCCCCCNCCNCCNCCNCCNCCNCC 1
RESULT 10
CR154670/c 50 bp DNA linear GSS 06-JUL-2004
LOCUS CR154670
DEFINITION Reverse strand read from insert in 3'HPRT insertion targeting and

```

```

chromosome engineering clone MHPP120c05, genomic survey sequence.
CR154670
ACCESSION
VERSION CR154670.1 GI:49933515
KEYWORDS GSS; genome survey sequence; MTCR.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 50)
ADAMS,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,
JONKERS,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,
Rogers,J. and Bradley,A.
Direct Submission
TITLE Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
JOURNAL CB10 1SA, UK. http://www.sanger.ac.uk/MTCR
FEATURES
source
1..50
Location/Qualifiers
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="MHPP120c05"
/clone_lib="MHPP"

ORIGIN
Alignment Scores:
Pred. No.: 4.55e+04 Length: 50
Score: 65.00 Matches: 12
Percent Similarity: 42.86% Conservative: 0
Best Local Similarity: 42.86% Mismatches: 4
Query Match: 3.40% Indels: 12
DB: 9 Gaps: 1

US-09-544-776-2 (1-373) x CR154670 (1-50)
QY 143 ProProProProProAlaSerValSerProGlnAlaGluProValTrpThrProPro 162
Db 50 CCCCCCCCCCCCCCCCCC-----CCCCC 27

QY 163 AlaProAlaProAlaAlaProPro 170
Db 26 CCCCCCNCCCCCCCCCCCCCCCCC 3

RESULT 11
AZ646447 48 bp DNA linear GSS 14-DEC-2000
LOCUS IM0512A06R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
DEFINITION clone UUGC1M0512A06 R, genomic survey sequence.
ACCESSION AZ646447
VERSION AZ646447.1 GI:11776922
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 48)
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,
Kelly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von
Niederhausern,A. and Wright,D.,Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
TITLE Unpublished (2000)
JOURNAL Contact: Robert B. Weiss
COMMENT University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0512 row: A column: 06
Seq primer: CACACAGAAACAGCTATGACC

```

```

Class: plasmid ends
High quality sequence stop: 48.
FEATURES
source
1..48
Location/Qualifiers
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="UUGC1M0512A06"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, TI-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

ORIGIN
Alignment Scores:
Pred. No.: 5.03e+04 Length: 48
Score: 64.00 Matches: 12
Percent Similarity: 42.86% Conservative: 0
Best Local Similarity: 42.86% Mismatches: 4
Query Match: 3.35% Indels: 12
DB: 8 Gaps: 1

US-09-544-776-2 (1-373) x AZ646447 (1-48)
QY 143 ProProProProProAlaSerValSerProGlnAlaGluProValTrpThrProPro 162
Db 1 CCCCCCCCCCCCCCCCCC-----CCCCC 24

QY 163 AlaProAlaProAlaAlaProPro 170
Db 25 CCCCCCCCCCCCCCCCCCCCCC 48

RESULT 12
AZ506149 49 bp DNA linear GSS 05-OCT-2000
LOCUS IM0347J05F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
DEFINITION clone UUGC1M0347J05 F, genomic survey sequence.
ACCESSION AZ506149
VERSION AZ506149.1 GI:10687465
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 49)
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,
Kelly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von
Niederhausern,A. and Wright,D.,Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
TITLE Unpublished (2000)
JOURNAL Contact: Robert B. Weiss
COMMENT University of Utah Genome Center

```

University of Utah  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT  
84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: ddunn@genetics.utah.edu  
Insert Length: 10000 Std Error: 0.00  
Plate: 0347 row: J column: 05  
Seq primer: CGTTGTAAACGACGCGCCAGT  
Class: plasmid ends  
High quality sequence stop: 49.

## FEATURES

source

1. 49  
Location/Qualifiers  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="UUGC1M0347005"  
/sex="Male"  
/lab\_host="E. Coli strain XL10-Gold, Tl-resistant, F-"  
/clone\_lib="Mouse 10kb plasmid UUGC1M library"  
/notes="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (GI|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

## ORIGIN

Alignment Scores:  
Pred. No.: 5.16e+04 Length: 49  
Score: 64.00 Matches: 12  
Percent Similarity: 42.86% Conservative: 0  
Best Local Similarity: 42.86% Mismatches: 4  
Query Match: 3.35% Indels: 12  
DB: 8 Gaps: 1

US-09-544-776-2 (1-373) x AZ506149 (1-49)

QY 143 ProProProProProAlaSerValSerProGlnAlaGluProValTrpThrProPro 162  
Db 2 CCCCCCCCCCCCCCCCCC-----CCCCC 25  
QY 163 AlaProAlaProAlaAlaProPro 170  
Db 26 CCCCCCCCCCCCCCTCCCCCCCC 49

## RESULT 13

AZ656875

LOCUS

DEFINITION AZ656875 49 bp DNA linear GSS 14-DEC-2000  
1M0532M11R Mouse 10kb plasmid UUGC1M library Mus musculus genomic  
clone UUGC1M0532M11 R, genomic survey sequence.

ACCESSION

AZ656875.1 GI:11794021

VERSION

GSS.

KEYWORDS

SOURCE Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 49)

## AUTHORS

Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,  
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T.,  
Reilly, M., Rose, R., Rose, R., Stokes, R., Tingey, A., von  
Niederhausern, A. and Wright, D., Weiss, R.  
Mouse whole genome scaffolding with paired end reads from 10kb  
plasmid inserts

## JOURNAL

Unpublished (2000)

## COMMENT

Contact: Robert B. Weiss  
University of Utah  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT  
84112, USA

Tel: 801 585 5606  
Fax: 801 585 7177  
Email: ddunn@genetics.utah.edu  
Insert Length: 10000 Std Error: 0.00  
Plate: 0532 row: M column: 11  
Seq primer: CACACAGGAACAGCTATGACC  
Class: plasmid ends  
High quality sequence stop: 49.

## FEATURES

source

1. 49  
Location/Qualifiers  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="UUGC1M0532M11"  
/sex="Male"  
/lab\_host="E. Coli strain XL10-Gold, Tl-resistant, F-"  
/clone\_lib="Mouse 10kb plasmid UUGC1M library"  
/notes="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (GI|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

## ORIGIN

Alignment Scores:  
Pred. No.: 5.16e+04 Length: 49  
Score: 64.00 Matches: 12  
Percent Similarity: 42.86% Conservative: 0  
Best Local Similarity: 42.86% Mismatches: 4  
Query Match: 3.35% Indels: 12  
DB: 8 Gaps: 1

US-09-544-776-2 (1-373) x AZ656875 (1-49)

QY 143 ProProProProProAlaSerValSerProGlnAlaGluProValTrpThrProPro 162  
Db 1 CCCCCCCCCCCCCCCCCC-----CCCCC 24

QY 163 AlaProAlaProAlaAlaProPro 170

Db 25 CCCCCCCCCCCCCCCCCCCCCC 48

## RESULT 14

AZ764533

LOCUS

DEFINITION AZ764533 49 bp DNA linear GSS 16-FEB-2001  
1M0560017R Mouse 10kb plasmid UUGC1M library Mus musculus genomic

clone UUGC1M0560017 R, genomic survey sequence.	clone UUGC1M0584E23 R, genomic survey sequence.
ACCESSION AZ764533	ACCESSION AZ773388
VERSION AZ764533.1	VERSION AZ773388.1
KEYWORDS GSS.	KEYWORDS GSS.
SOURCE Mus musculus (house mouse)	SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus	ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 49)	REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 49)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,R., Stokes,R., Tingley,A., von Niederhausern,A. and Wright,D., Weiss,R.	AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,R., Stokes,R., Tingley,A., von Niederhausern,A. and Wright,D., Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts	TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished (2000)	JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss University of Utah Genome Center University of Utah Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA Tel: 801 585 5606 Fax: 801 585 7177 Email: ddunn@genetics.utah.edu Insert Length: 10000 Std Error: 0.00 Plate: 0560 row: 0 column: 17 Seq primer: CACACAGGAACAGCTATGACC Class: plasmid ends High quality sequence stop: 49.	COMMENT Contact: Robert B. Weiss University of Utah Genome Center University of Utah Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA Tel: 801 585 5606 Fax: 801 585 7177 Email: ddunn@genetics.utah.edu Insert Length: 10000 Std Error: 0.00 Plate: 0584 row: E column: 23 Seq primer: CACACAGGAACAGCTATGACC Class: plasmid ends High quality sequence stop: 49.
FEATURES source 1..49 /organism="Mus musculus" /mol_type="genomic DNA" /strain="C57BL/6J" /db_xref="taxon:10090" /clone="UUGC1M0560017" /sex="Male" /lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-" /clone_lib="Mouse 10kb plasmid UUGC1M library" /note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi 4732114 gb AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."	FEATURES source 1..49 /organism="Mus musculus" /mol_type="genomic DNA" /strain="C57BL/6J" /db_xref="taxon:10090" /clone="UUGC1M0584E23" /sex="Male" /lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-" /clone_lib="Mouse 10kb plasmid UUGC1M library" /note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi 4732114 gb AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
ORIGIN Alignment Scores: Pred. No.: Score: Percent Similarity:	ORIGIN Alignment Scores: Pred. No.: Score: Percent Similarity:
5.16e+04 64.00 42.86%	5.16e+04 64.00 42.86%
13 ProProProProProProAlaSerValSerProGlnAlaGluProValTrpThrProPro 162 	13 ProProProProProProAlaSerValSerProGlnAlaGluProValTrpThrProPro 162 
1 CCCCCCCCCCCCCCCCCC-----CCCCC 24	1 CCCCCCCCCCCCCCCCCC-----CCCCC 24

Best Local Similarity: 42.86%  
 Query Match: 3.35%  
 DB: 8  
 Mismatches: 4  
 Indels: 12  
 Gaps: 1

US-09-544-776-2 (1-373) x AZ773388 (1-49)

Qy	143	ProProProProProProAlaSerValSerProGlnAlaGluProValTrpThrProPro	162
Db	1	CCCCCCCCCCCCCCCCCCCC	CCCCC 24
Qy	163	AlaProAlaProAlaProPro	170
Db	25	CCCCCCCCCCCCCCCCCCCC	CCCCC 48

Search completed: August 4, 2005, 02:03:21  
 Job time : 3091 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 3, 2005, 17:14:22 ; Search time 206 Seconds  
(without alignments)  
2962.773 Million cell updates/sec

Title: US-09-544-776-2  
Perfect score: 1910  
Sequence: 1 MEDLOSPLVSSSDSPRPQ.....VKDAMAKIQAKIPGLKRAE 373

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 1209694

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp  
-Q=/cgn2\_1/USPTO\_spool\_p/US09544776/runat\_03082005.181415.27646/app.query.fasta\_1.519  
-DB=issued\_Patents\_NA -QFMT=fastap -SUFFIX=p2nslm50.rni -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZ=500 -MINLEN=0  
-MAXLEN=50 -USER=US09544776 @CEN 1.1.105 @runat 03082005.181415.27646 -NCFU=6  
-ICPU=3 -NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA.\*  
1: /cgn2\_6/ptodata/1/ina/5A COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PCTUS COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	68	3.6	48	3	US-08-979-608A-36
2	68	3.6	48	4	Sequence 36, Appl
3	68	3.6	48	4	US-09-517-849-36
4	66	3.5	48	3	Sequence 36, Appl
5	66	3.5	48	3	US-09-580-923-34
6	66	3.5	50	3	Sequence 34, Appl
7	63	3.3	50	3	US-08-860-038-17
8	63	3.3	50	3	Sequence 17, Appl
9	63	3.3	50	3	US-09-580-923-17
10	59	3.1	50	3	Sequence 17, Appl
11	59	3.1	50	3	US-08-956-171E-2153
12	59	3.1	50	3	Sequence 2153, Ap
					US-08-781-986A-2153
					Sequence 14, Appl
					US-09-930-181-14
					Sequence 14, Appl
					US-08-979-608A-30
					Sequence 30, Appl
					US-09-517-849-30
					Sequence 30, Appl
					US-09-616-289-30
					Sequence 30, Appl

13	58	3.0	38	1	US-08-563-864-1	Sequence 1, Appl
14	58	3.0	38	5	PCT-US95-06833-1	Sequence 1, Appl
C 15	57	3.0	39	1	US-08-068-747-9	Sequence 9, Appl
C 16	57	3.0	45	4	US-09-590-211A-9	Sequence 9, Appl
C 17	57	3.0	48	3	US-09-580-923-34	Sequence 34, Appl
C 18	57	3.0	48	4	US-09-580-923-34	Sequence 34, Appl
19	57	3.0	48	4	US-09-119-507B-78	Sequence 78, Appl
20	57	3.0	48	4	US-08-897-556A-78	Sequence 78, Appl
C 21	57	3.0	48	4	US-09-547-693-78	Sequence 78, Appl
C 22	57	3.0	50	3	US-08-860-038-17	Sequence 17, Appl
C 23	56	2.9	48	4	US-09-580-923-17	Sequence 17, Appl
24	56	2.9	48	4	US-09-119-507B-88	Sequence 88, Appl
25	56	2.9	48	4	US-08-897-556A-88	Sequence 88, Appl
26	53	2.8	30	1	US-08-068-747-2	Sequence 88, Appl
C 27	53	2.8	30	1	US-08-068-747-2	Sequence 2, Appl
C 28	53	2.8	30	1	US-08-068-747-3	Sequence 3, Appl
C 29	53	2.8	30	3	US-09-083-123-2	Sequence 4, Appl
C 30	53	2.8	30	3	US-09-083-123-2	Sequence 2, Appl
C 31	53	2.8	30	3	US-09-083-123-8	Sequence 8, Appl
C 32	53	2.8	30	3	US-08-589-109A-12	Sequence 12, Appl
C 33	53	2.8	31	3	US-09-083-123-4	Sequence 11, Appl
C 34	53	2.8	32	3	US-09-196-270-11	Sequence 5, Appl
C 35	53	2.8	33	1	US-09-083-123-5	Sequence 7, Appl
C 36	53	2.8	33	4	US-08-068-747-7	Sequence 7, Appl
C 37	53	2.8	35	4	US-09-590-211A-5	Sequence 5, Appl
C 38	53	2.8	35	4	US-09-196-270-10	Sequence 10, Appl
C 39	53	2.8	36	1	US-08-068-747-8	Sequence 8, Appl
C 40	53	2.8	36	4	US-09-590-211A-6	Sequence 6, Appl
C 41	53	2.8	39	1	US-08-068-747-9	Sequence 9, Appl
C 42	53	2.8	39	4	US-09-590-211A-7	Sequence 7, Appl
C 43	53	2.8	40	3	US-09-476-256-29	Sequence 29, Appl
C 44	53	2.8	42	4	US-09-590-211A-8	Sequence 8, Appl
C 45	53	2.8	43	3	US-09-411-722-4	Sequence 4, Appl
C 45	53	2.8	43	4	US-09-855-266A-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1

US-08-979-608A-36  
; Sequence 36, Application US/08979608A  
; Patent No. 6355451

GENERAL INFORMATION:

APPLICANT: Lees, Ann M.

Law, Robert S.

Arjona, Anibal A.

TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN

BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND

TREATING ATHEROSCLEROSIS

NUMBER OF SEQUENCES: 42

CORRESPONDENCE ADDRESS:

ADDRESSER: Fish & Richardson P.C.

STREET: 225 Franklin Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02110-2804

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION NUMBER: US/08/979,608A

APPLICATION DATA:

FILING DATE: 26-NOV-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/048,547

FILING DATE: 03-JUN-1997

APPLICATION NUMBER: US 60/031,930

FILING DATE: 27-NOV-1996

ATTORNEY/AGENT INFORMATION:

NAME: Myers, Louis

REGISTRATION NUMBER: 35,965

```

; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 1...48
; SEQUENCE DESCRIPTION: SEQ ID NO: 36:
US-09-517-849-36

Alignment Scores:
Pred. No.:      892                               Length:      48
Score:          68.00                             Matches:       11
Percent Similarity: 100.00%                       Conservative:   5
Best Local Similarity: 68.75%                      Mismatches:    0
Query Match:     3.56%                              Indels:        0
DB:              4                                 Gaps:         0

US-09-544-776-2 (1-373) x US-09-517-849-36 (1-48)

Qy      33 AspGluGluGluGluGluGluGluGluGluGluAspGluAspGluasp 48
      ::::|||||
Db      1 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGCAGCAGGACGACGACGCAC 48

RESULT 3
US-09-616-289-36
; Sequence 36, Application US/09616289
; Patent No. 6632923
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/616,289
; CURRENT FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 36
; LENGTH: 48
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-616-289-36

Alignment Scores:
Pred. No.:      892                               Length:      48
Score:          68.00                             Matches:       11
Percent Similarity: 100.00%                       Conservative:   5
Best Local Similarity: 68.75%                      Mismatches:    0
Query Match:     3.56%                              Indels:        0
DB:              4                                 Gaps:         0

US-09-544-776-2 (1-373) x US-09-616-289-36 (1-48)

Qy      33 AspGluGluGluGluGluGluGluGluGluGluAspGluAspGluasp 48
      ::::|||||
Db      1 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGCAGCAGGACGACGACGCAC 48

RESULT 4
US-09-580-923-34
; Sequence 34, Application US/09580923
; Patent No. 6319672
; GENERAL INFORMATION:
; APPLICANT: Crouzet, Joel
; APPLICANT: Scherman, Daniel
```

```

; REFERENCE/DOCKET NUMBER: ST94090-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3816
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 50 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "Oligonucleotide"
US-08-860-038-17

Alignment Scores:
Pred. No.: 1.29e+03 Length: 50
Score: 66.00 Matches: 12
Percent Similarity: 100.00% Conservative: 3
Best Local Similarity: 80.00% Mismatches: 0
Query Match: 3.46% Indels: 0
DB: 3 Gaps: 0

US-09-544-776-2 (1-373) x US-08-860-038-17 (1-50)

Qy 32 GluAspGluGluGluGluGluGluGluGluGluGluAspGluAsp 46
Db 4 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 48

RESULT 6
US-09-580-923-17
; Sequence 17, Application US/09580923
; Patent No. 6319672
; GENERAL INFORMATION:
; APPLICANT: Crouzet, Joel
; APPLICANT: Scherman, Daniel
; APPLICANT: Wils, Pierre
; APPLICANT: Cameron, Beatrice
; APPLICANT: Blanche, Francis
; TITLE OF INVENTION: PURIFICATION OF A TRIPLE HELIX FORMATION WITH AN
; TITLE OF INVENTION: IMMOBILIZED OLIGONUCLEOTIDE
; FILE REFERENCE: 03804.0138-01
; CURRENT APPLICATION NUMBER: US/09/580,923
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 08/860,038
; PRIOR FILING DATE: 1997-06-09
; PRIOR APPLICATION NUMBER: PCT/FR95/01468
; PRIOR FILING DATE: 1995-11-08
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
; LENGTH: 50
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: oligonucleotide
US-09-580-923-17

Alignment Scores:
Pred. No.: 1.29e+03 Length: 50
Score: 66.00 Matches: 12
Percent Similarity: 100.00% Conservative: 3
Best Local Similarity: 80.00% Mismatches: 0
Query Match: 3.46% Indels: 0
DB: 3 Gaps: 0

US-09-544-776-2 (1-373) x US-09-580-923-17 (1-50)

Qy 32 GluAspGluGluGluGluGluGluGluGluGluGluAspGluAsp 46
Db 4 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 48

RESULT 7

```



```

Pred. No.: 3.36e+03 Length: 50
Score: 60.00 Matches: 12
Percent Similarity: 46.11% Conservative: 0
Best Local Similarity: 46.11% Mismatches: 2
Query Match: 3.14% Indels: 12
DB: 3 Gaps: 1

US-09-544-776-2 (1-373) x US-09-930-181-14 (1-50)

Qy 143 PROPROProProProAlaSerValSerProGlnAlaGluProValTrpThrProPro 162
Db 42 CGCGCGCGCGCGCGCGCG-----CGCGCG 19

Qy 163 AlaProAlaProAlaAla 168
Db 18 AGCGCGCGCGCGCGCGCGC 1

RESULT 10
US-08-979-608A-30
; Sequence 30, Application US/08979608A
; Patent No. 6355451
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; Lees, Robert S.
; Law, Simon W.
; Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
; TREATING ATHEROSCLEROSIS
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/979,608A
; FILING DATE: 26-No. 6355451-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/048,547
; FILING DATE: 03-JUN-1997
; APPLICATION NUMBER: US 60/031,930
; FILING DATE: 27-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 45 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 1...45
; SEQUENCE DESCRIPTION: SEQ ID NO: 30:
US-08-979-608A-30

Alignment Scores:
Pred. No.: 3.43e+03 Length: 45
Score: 59.00 Matches: 9
Percent Similarity: 100.00% Conservative: 6

```



MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
PCT-US95-06839-1

Alignment Scores:  
Pred. No.: 3.22e+03 Length: 38  
Score: 58.00 Matches: 9  
Percent Similarity: 75.00% Conservative: 0  
Best Local Similarity: 75.00% Mismatches: 3  
Query Match: 3.04% Indels: 0  
DB: 5 Gaps: 0

US-09-544-776-2 (1-373) x PCT-US95-06839-1 (1-38)

QY 137 AspGluProAlaArgProProProPro 148  
Db 1 GATCCACCCACCACCACCACCACCACCT 36

RESULT 15  
US-08-068-747-9/c  
Sequence 9, Application US/08068747  
Patent No. 5695933  
GENERAL INFORMATION:  
APPLICANT: Schalling, Martin  
APPLICANT: Hudson, Thomas J.  
APPLICANT: Housman, David E.  
TITLE OF INVENTION: Direct Determination of Expanded  
TITLE OF INVENTION: Nucleotide Repeats in the Human Genome  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
STREET: Two Militia Drive  
CITY: Lexington  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02173

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/068,747  
FILING DATE: 28-MAY-1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Granahan, Patricia  
REGISTRATION NUMBER: 32,227  
REFERENCE/DOCKET NUMBER: MIT-6141  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-861-6240  
TELEFAX: 617-861-9540  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 39 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: other nucleic acid  
DESCRIPTION: /desc = "Synthetic"

Alignment Scores:  
Pred. No.: 3.91e+03 Length: 39  
Score: 57.00 Matches: 11  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 91.67% Mismatches: 0  
Query Match: 2.98% Indels: 0  
DB: 1 Gaps: 0

US-09-544-776-2 (1-373) x US-08-068-747-9 (1-39)

QY 32 GluAspGluGluGluGluGluGluGluGluGlu 43  
Db 37 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2

Search completed: August 3, 2005, 23:49:49  
Job time : 208 secs

**THIS PAGE BLANK (05/10)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 3, 2005, 17:14:22 ; Search time 686 Seconds  
(without alignments)  
3519.226 Million cell updates/sec

Title: US-09-544-776-2

Perfect score: 1910

Sequence: 1 MEDDQSPVSSSDSPRPQ.....VKDAMAKIOAKIPGLKRAE 373

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 7287783 seqs, 3236178273 residues

Total number of hits satisfying chosen parameters: 8279544

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp  
-Q=/cgn2\_1/USPTO\_spool\_p/US09544776/runat\_03082005\_181415\_27663/app\_query.fasta\_1.519  
-DB=Published Applications NA -QFMT=fastap -SURFIX=p2nszlm50.rnpb  
-MINMATCH=0.1 -LOOPCL=0 -LOPEXT=0 -UNITS=bits -START=1 -END=-1  
-MATRIX=biosum62 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct  
-THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext  
-HEAFSIZE=500 -MINLEN=0 -MAXLEN=50  
-USER=US09544776 -CGEN 1.1 723 @runat\_03082005\_181415\_27663 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/2/pubpna/US10D\_PUBCOMB.seq.\*  
17: /cgn2\_6/ptodata/2/pubpna/US10E\_PUBCOMB.seq.\*  
18: /cgn2\_6/ptodata/2/pubpna/US10F\_PUBCOMB.seq.\*  
19: /cgn2\_6/ptodata/2/pubpna/US10G\_PUBCOMB.seq.\*  
20: /cgn2\_6/ptodata/2/pubpna/US10H\_PUBCOMB.seq.\*  
21: /cgn2\_6/ptodata/2/pubpna/US10I\_PUBCOMB.seq.\*  
22: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*  
23: /cgn2\_6/ptodata/2/pubpna/US11A\_PUBCOMB.seq.\*  
24: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq.\*  
25: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*  
26: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	68	3.6	48	9	US-09-962-055-36	Sequence 36, Appl
2	68	3.6	48	9	US-09-976-740-36	Sequence 36, Appl
3	68	3.6	48	13	US-10-023-529-36	Sequence 36, Appl
4	68	3.6	48	13	US-10-023-523-36	Sequence 36, Appl
5	68	3.6	48	17	US-10-616-187-36	Sequence 36, Appl
6	68	3.6	48	18	US-10-671-242-36	Sequence 36, Appl
7	66	3.5	48	16	US-10-275-071-34	Sequence 34, Appl
8	66	3.5	50	16	US-10-275-071-17	Sequence 17, Appl
9	63	3.3	50	8	US-08-781-986A-2153	Sequence 2153, Ap
10	63	3.3	50	18	US-10-329-624-2153	Sequence 2153, Ap
C 11	60	3.1	50	14	US-10-195-072-14	Sequence 14, Appl
C 12	60	3.1	50	14	US-10-195-071-14	Sequence 14, Appl
C 13	60	3.1	50	19	US-10-627-253A-369	Sequence 369, App
C 14	60	3.1	50	19	US-10-627-253A-370	Sequence 370, App
15	59	3.1	45	9	US-09-962-055-30	Sequence 30, Appl
16	59	3.1	45	9	US-09-976-740-30	Sequence 30, Appl
17	59	3.1	45	13	US-10-023-529-30	Sequence 30, Appl
18	59	3.1	45	13	US-10-023-523-30	Sequence 30, Appl
19	59	3.1	45	17	US-10-616-187-30	Sequence 30, Appl
20	59	3.1	45	18	US-10-671-242-30	Sequence 30, Appl
C 21	58	3.0	34	21	US-10-455-453-28	Sequence 28, Appl
C 22	58	3.0	40	21	US-10-455-453-29	Sequence 29, Appl
C 23	57	3.0	36	17	US-10-418-182-86	Sequence 86, Appl
C 24	57	3.0	45	21	US-10-889-263-9	Sequence 9, Appl
C 25	57	3.0	45	21	US-10-831-819-14	Sequence 14, Appl
C 26	57	3.0	47	19	US-10-627-253A-373	Sequence 373, App
C 27	57	3.0	47	19	US-10-627-253A-374	Sequence 374, App
C 28	57	3.0	48	16	US-10-275-071-34	Sequence 34, Appl
C 29	57	3.0	48	17	US-10-437-708-78	Sequence 78, Appl
C 30	57	3.0	48	17	US-10-395-402-78	Sequence 78, Appl
C 31	57	3.0	48	20	US-10-257-199-78	Sequence 78, Appl
C 32	57	3.0	48	21	US-10-418-032-78	Sequence 78, Appl
C 33	57	3.0	50	16	US-10-275-071-17	Sequence 17, Appl
C 34	56	2.9	42	9	US-09-750-021-5	Sequence 5, Appl
C 35	56	2.9	42	17	US-10-437-708-88	Sequence 88, Appl
C 36	56	2.9	48	17	US-10-395-402-88	Sequence 88, Appl
C 37	56	2.9	48	20	US-10-257-199-88	Sequence 88, Appl
C 38	56	2.9	48	21	US-10-418-032-88	Sequence 88, Appl
C 39	56	2.9	50	18	US-10-062-188-96	Sequence 96, Appl
C 40	55	2.9	38	19	US-10-627-253A-377	Sequence 377, App
C 41	55	2.9	38	19	US-10-627-253A-378	Sequence 378, App
C 42	54	2.8	36	19	US-10-296-085A-74	Sequence 74, Appl
C 43	54	2.8	42	10	US-09-833-203-8	Sequence 8, Appl
C 44	54	2.8	42	10	US-09-833-203-9	Sequence 9, Appl
C 45	54	2.8	42	15	US-10-211-296-11	Sequence 11, Appl

# ALIGNMENTS

## RESULT 1

US-09-962-055-36  
; Sequence 36, Application US/09962055  
; Patent No. US20020052033A1  
; GENERAL INFORMATION:

APPLICANT: Lees, Ann M.  
; Lees, Robert S.  
; Law, Simon W.  
; Artjona, Anibal A.

TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN

BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND  
TREATING ATHEROSCLEROSIS

NUMBER OF SEQUENCES: 42

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson P.C.







; TITLE OF INVENTION: Full-Length Serine Protein Kinase in Brain and Pancreas	
; FILE REFERENCE: 16U 101 C2	
; CURRENT APPLICATION NUMBER: US/10/195,072	
; CURRENT FILING DATE: 2002-07-15	
; PRIOR APPLICATION NUMBER: US 09/930,181	
; PRIOR FILING DATE: 2001-08-16	
; NUMBER OF SEQ ID NOS: 18	
; SOFTWARE: PatentIn version 3.1	
; SEQ ID NO 14	
; LENGTH: 50	
; TYPE: DNA	
; ORGANISM: Homo sapiens	
US-10-195-072-14	
Alignment Scores:	
Pred. No.:	2.37e+03
Score:	60.00
Percent Similarity:	46.15%
Best Local Similarity:	46.15%
Query Match:	3.14%
DB:	14
US-09-544-776-2 (1-373) x US-10-195-072-14 (1-50)	
Qy	143 ProProProProProProAlaSerProGlnAlaGluProValTrpThrProPro 162
Db	42 CCGCGCGCGCGCGCGCG-----CGCGCG 19
Qy	163 AlaProAlaProAlaAla 168
Db	18 AGGCGCGCGCGCGCGCC 1
RESULT 12	
US-10-195-071-14/c	
; Sequence 14, Application US/10195071	
; Publication No. US20030096271A1	
; GENERAL INFORMATION:	
; APPLICANT: Origene Technologies	
; TITLE OF INVENTION: Full-Length Serine Protein Kinase in Brain and Pancreas	
; FILE REFERENCE: 16U 101 C1	
; CURRENT APPLICATION NUMBER: US/10/195,071	
; CURRENT FILING DATE: 2002-07-15	
; PRIOR APPLICATION NUMBER: US 09/930,181	
; PRIOR FILING DATE: 2001-08-16	
; NUMBER OF SEQ ID NOS: 18	
; SOFTWARE: PatentIn version 3.1	
; SEQ ID NO 14	
; LENGTH: 50	
; TYPE: DNA	
; ORGANISM: Homo sapiens	
US-10-195-071-14	
Alignment Scores:	
Pred. No.:	2.37e+03
Score:	60.00
Percent Similarity:	46.15%
Best Local Similarity:	46.15%
Query Match:	3.14%
DB:	14
US-09-544-776-2 (1-373) x US-10-195-071-14 (1-50)	
Qy	143 ProProProProProProAlaSerProGlnAlaGluProValTrpThrProPro 162
Db	42 CCGCGCGCGCGCGCGCG-----CGCGCG 19
Qy	163 AlaProAlaProAlaAla 168
Db	18 AGGCGCGCGCGCGCGCC 1
RESULT 13	
US-10-627-253A-369	
; Sequence 369, Application US/10627253A	

```
; Publication No. US20040161768A1
; GENERAL INFORMATION:
; APPLICANT: BRINKMANN, ULRICH
; APPLICANT: HOFFMEYER, SVEN
; APPLICANT: MORNHINWEG, ESTHER
; TITLE OF INVENTION: POLYMORPHISMS IN THE HUMAN GENE FOR THE MULTIDRUG
; TITLE OF INVENTION: RESISTANCE-ASSOCIATED PROTEIN 1 (MRP-1) AND THEIR USE IN
; TITLE OF INVENTION: DIAGNOSTIC AND THERAPEUTIC APPLICATIONS
; FILE REFERENCE: VOS-42 CON
; CURRENT APPLICATION NUMBER: US/10/627,253A
; PRIOR FILING DATE: 2003-07-24
; PRIOR APPLICATION NUMBER: PCT/EP02/00796
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: EP 01101651.6
; PRIOR FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 406
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 369
; LENGTH: 50
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide
US-10-627-253A-369

Alignment Scores:
Pred. No.:          2,37e+03      Length:       50
Score:              60.00         Matches:        13
Percent Similarity: 46.43%       Conservative:    0
Best Local Similarity: 46.43%     Mismatches:      3
Query Match:        3.14%        Indels:         12
DB:                 19           Gaps:            1

US-09-544-776-2 (1-373) x US-10-627-253A-369 (1-50)

Qy   140 ProAlaArgProProProProProAlaSerValSerProGlnAlaGluProValTrp 159
Db   3 CCTGCAGCGCCGCCGGCGCGCGCGCGC----- 29

Qy   160 ThrProAlaProAlaProAla 167
Db   30 ---CCGCCGCCGCCAGCGTAGCG 50

RESULT 14
US-10-627-253A-370/c
; Sequence 370, Application US/10627253A
; Publication No. US20040161768A1
; GENERAL INFORMATION:
; APPLICANT: BRINKMANN, ULRICH
; APPLICANT: HOFFMEYER, SVEN
; APPLICANT: MORNHINWEG, ESTHER
; TITLE OF INVENTION: POLYMORPHISMS IN THE HUMAN GENE FOR THE MULTIDRUG
; TITLE OF INVENTION: RESISTANCE-ASSOCIATED PROTEIN 1 (MRP-1) AND THEIR USE IN
; TITLE OF INVENTION: DIAGNOSTIC AND THERAPEUTIC APPLICATIONS
; FILE REFERENCE: VOS-42 CON
; CURRENT APPLICATION NUMBER: US/10/627,253A
; CURRENT FILING DATE: 2003-07-24
; PRIOR APPLICATION NUMBER: PCT/EP02/00796
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: EP 01101651.6
; PRIOR FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 406
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 370
; LENGTH: 50
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide
US-10-627-253A-370

Alignment Scores:
Pred. No.:          2,37e+03      Length:       50
```



**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 3, 2005, 17:13:53 ; Search time 703 Seconds  
(without alignments)  
3434.124 Million cell updates/sec

Title: US-09-544-776-2  
Perfect score: 1910  
Sequence: 1 MEDDQSLVSSSDSPRPQ.....VKDAMAKIQAKIFGLKRAE 373

Scoring table: BLOSUM62 Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 7287783 seqs, 3236178273 residues

Total number of hits satisfying chosen parameters: 14575566

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlp  
-Q=/cgn2\_1/USPTO.spool\_p/US09544776/runat\_03082005\_181346\_28256/app\_query.fasta\_1.519  
-DB=Published Applications NA -QFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEADSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -CGEN=US09544776 @CGEN 1 1 723 @runat\_03082005\_181346\_28256  
-NCPU=6 -ICPU=3 -NO\_MMAP -LARGESQUEW -NEG SCORES=0 -WAIT -DSBLOCk=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:  
1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:  
9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:  
13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:  
14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:  
15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq:  
16: /cgn2\_6/ptodata/2/pubpna/US10D\_PUBCOMB.seq:  
17: /cgn2\_6/ptodata/2/pubpna/US10E\_PUBCOMB.seq:  
18: /cgn2\_6/ptodata/2/pubpna/US10F\_PUBCOMB.seq:  
19: /cgn2\_6/ptodata/2/pubpna/US10G\_PUBCOMB.seq:  
20: /cgn2\_6/ptodata/2/pubpna/US10H\_PUBCOMB.seq:  
21: /cgn2\_6/ptodata/2/pubpna/US10I\_PUBCOMB.seq:  
22: /cgn2\_6/ptodata/2/pubpna/US11A\_PUBCOMB.seq:  
23: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq:  
24: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq:  
25: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:  
26: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1879	98.4	1122	9	US-09-789-386-5
2	1879	98.4	1610	9	US-09-765-205-5
3	1879	98.4	1610	21	US-10-347-669-5
4	1879	98.4	2050	21	US-10-956-157-4530
5	1879	98.4	2226	21	US-10-956-157-4531
6	1879	98.4	2235	14	US-10-060-036-54
7	1871	98.0	2052	14	US-10-466-258-3
8	1871	98.0	2052	22	US-10-466-391A-3
9	1566	82.0	2782	15	US-10-205-194-165
10	1459.5	76.4	3579	9	US-09-789-386-1
11	1459.5	76.4	3579	9	US-09-893-348-22
12	1459.5	76.4	3579	18	US-10-267-502-212
13	1459.5	76.4	3579	19	US-10-327-213-8
14	1459.5	76.4	3579	19	US-10-466-258-8
15	1459.5	76.4	3579	20	US-10-810-653-22
16	1459.5	76.4	3579	22	US-10-466-391A-8
17	1459.5	76.4	4053	9	US-09-758-140-5
18	1459.5	76.4	4053	9	US-09-972-599A-5
19	1459.5	76.4	4053	19	US-10-717-597-310
20	1459.5	76.4	4623	21	US-10-956-157-4532
21	1459.5	76.4	4632	14	US-10-060-036-53
22	1173	61.4	4684	9	US-09-893-348-17
23	1173	61.4	4684	20	US-10-810-653-17
24	1133	59.3	3492	18	US-10-267-502-214
25	1018.5	53.3	2883	21	US-10-956-157-4529
26	942	49.3	3478	21	US-10-956-157-4528
27	931	48.7	1400	21	US-10-956-157-9763
28	927	48.5	1980	17	US-10-220-891-22
29	918	48.1	600	21	US-10-956-157-9764
30	917	48.0	799	18	US-10-660-946-2
31	917	48.0	1160	14	US-10-175-523-156
32	917	48.0	1609	21	US-10-956-157-4527
33	917	48.0	1785	17	US-10-439-388-62
34	917	48.0	1785	21	US-10-956-157-1705
35	908	47.5	994	11	US-09-978-360A-110
36	904	47.3	2610	18	US-10-641-643-382
37	876	45.9	1798	19	US-10-466-258-10
38	876	45.9	1798	22	US-10-466-391A-10
39	862	45.1	1400	21	US-10-956-157-9765
40	857	44.9	1514	9	US-09-823-245A-349
41	795	41.6	1400	21	US-10-956-157-9762
42	795	41.6	1400	21	US-10-956-157-9767
43	747.5	39.1	2331	18	US-10-267-502-213
44	747.5	39.1	3202	9	US-09-954-456-210
45	747.5	39.1	3202	17	US-10-172-118-386

ALIGNMENTS

RESULT 1  
US-09-789-386-5  
; Sequence 5, Application US/09789386  
; Patent No. US20020010324A1  
; GENERAL INFORMATION:  
; APPLICANT: MICHALOVICH, DAVID  
; APPLICANT: PRINJHA, RABINDER KUMAR  
; TITLE OF INVENTION: NOVEL COMPOUNDS  
; FILE REFERENCE: GP-30165-C1  
; CURRENT APPLICATION NUMBER: US/09/789,386  
; CURRENT FILING DATE: 2001-02-21  
; PRIOR APPLICATION NUMBER: U.K. 9916898.1  
; PRIOR FILING DATE: 1999-07-19  
; PRIOR APPLICATION NUMBER: U.K. 9816024.5  
; PRIOR FILING DATE: 1998-07-22  
; PRIOR APPLICATION NUMBER: US 09/359,208

```
; PRIOR FILING DATE: 1999-07-22
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 1122
; TYPE: DNA
; ORGANISM: HOMO SAPIENS
US-09-789-386-5

Alignment Scores:
Pred. No.: 6,29e-159 Length: 1122
Score: 1879.00 Matches: 370
Percent Similarity: 99.20% Conservatives: 0
Best Local Similarity: 99.20% Mismatches: 3
Query Match: 98.38% Indels: 0
DB: 9 Gaps: 0

US-09-544-776-2 (1-373) x US-09-789-386-5 (1-1122)

QY 1 MetGluAspLeuAspGlnSerProLeuValSerSerSerSerProProArgProGln 20
DB 1 ATGGAAGACTGGACCAAGTCTCTCTGGTCTCGTCTCGGACAGCCACCCCGGCGCAG 60

QY 21 ProAlaPheLysTyrGlnPheValArgGluProGluAspGluGluGluGlu 40
DB 61 CCGCGCTTCAAGTACCAGTTCTGTGGAGAGCCCGAGGACGAGGAGAGGAGGAG 120

QY 41 GluGluGluAspGluAspGluAspGluGluGluGluValLeuGluArgLysProAla 60
DB 121 GAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 180

QY 61 AlaGlyLeuSerAlaAlaProValProThrAlaProAlaAlaGlyAlaProLeuMetAsp 80
DB 181 GCCGGGCTGTCCGGGCGCCAGTGCACCGCCCTGCCCGCGCGCGCCCTCGATGGAC 240

QY 81 PheGlyAsnAspPheValProProAlaProAlaProAlaProAlaProVal 100
DB 241 TTCGGAATGACTTCTGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300

QY 101 AlaProGluArgGlnProSerTrpAspProSerProValSerSerThrValProAlaPro 120
DB 301 GCCCGGAGCGGAGCGCGCTTGGGACCGGAGCGCGGTGTCTGCGACCGCGCGCGCA 360

QY 121 SerPheLeuSerAlaAlaValSerProSerLysLeuProGluAspGluProPro 140
DB 361 TCCCCCGTGTCTCGCGGAGTCTCGCTTCCAGCTCCCTGAGGACGAGCGCTCCG 420

QY 141 AlaArgProProProProProProAlaSerValSerProGlnAlaGluProValTrpThr 160
DB 421 GCCCGGCTCCCTCTCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480

QY 161 ProProAlaProAlaProAlaProProSerThrProAlaAlaProLysArgArgGly 180
DB 481 CCGCCAGCCCGGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 540

QY 181 SerSerGlySerValValValAspLeuLeuTyrTrpArgAspLysLysThrGlyVal 200
DB 541 TCTCTGGGCTCAGTGTGTGTGACCTCTCTGAGGAGACATTAGGAGACTGGAGTG 600

QY 201 ValPheGlyAlaSerLeuPheLeuLeuLeuSerLeuThrValPheSerIleValSerVal 220
DB 601 GTGTTTGTGGCCAGCTATTCTCTGCTCTTTCATTGACAGTATTGACGATTGTGACGTA 660

QY 221 ThrAlaTyrIleAlaLeuAlaLeuLeuSerValThrIleSerProArgIleTyrLysGly 240
DB 661 ACAGCCTACATTCCCTTGGCCCTGCTCTCTGTGACCATCAGCTTTTAGGATATACAAGGT 720

QY 241 ValIleGlnAlaIleGlnLysSerAspGluGlyHisProPheArgAlaTyrLeuGluSer 260
DB 721 GTGATCCAGCTATCCAGAAATCAGATGAAGGCCACCCATTGAGGCAATATCTGGAATCT 780

QY 261 GluValAlaIleSerGluGluLeuValGlnLysTyrSerAsnSerAlaLeuGlyHisVal 280
DB 921 GTGATCCAGCTATCCAGAAATCAGATGAAGGCCACCCATTGAGGCAATATCTGGAATCT 980

; PRIOR FILING DATE: 1999-07-22
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 1122
; TYPE: DNA
; ORGANISM: HOMO SAPIENS
US-09-789-386-5

Alignment Scores:
Pred. No.: 6,29e-159 Length: 1122
Score: 1879.00 Matches: 370
Percent Similarity: 99.20% Conservatives: 0
Best Local Similarity: 99.20% Mismatches: 3
Query Match: 98.38% Indels: 0
DB: 9 Gaps: 0

US-09-544-776-2 (1-373) x US-09-789-386-5 (1-1122)

QY 1 MetGluAspLeuAspGlnSerProLeuValSerSerSerSerProProArgProGln 20
DB 1 ATGGAAGACTGGACCAAGTCTCTCTGGTCTCGTCTCGGACAGCCACCCCGGCGCAG 60

QY 21 ProAlaPheLysTyrGlnPheValArgGluProGluAspGluGluGluGluGlu 40
DB 61 CCGCGCTTCAAGTACCAGTTCTGTGGAGAGCCCGAGGACGAGGAGAGGAGGAG 120

QY 41 GluGluGluAspGluAspGluAspGluGluGluGluValLeuGluArgLysProAla 60
DB 121 GAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 180

QY 61 AlaGlyLeuSerAlaAlaProValProThrAlaProAlaAlaGlyAlaProLeuMetAsp 80
DB 181 GCCGGGCTGTCCGGGCGCCAGTGCACCGCCCTGCCCGCGCGCGCCCTCGATGGAC 240

QY 81 PheGlyAsnAspPheValProProAlaProAlaProAlaProAlaProVal 100
DB 241 TTCGGAATGACTTCTGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300

QY 101 AlaProGluArgGlnProSerTrpAspProSerProValSerSerThrValProAlaPro 120
DB 301 GCCCGGAGCGGAGCGCGCTTGGGACCGGAGCGCGGTGTCTGCGACCGCGCGCGCA 360

QY 121 SerPheLeuSerAlaAlaValSerProSerLysLeuProGluAspGluProPro 140
DB 361 TCCCCCGTGTCTCGCGGAGTCTCGCTTCCAGCTCCCTGAGGACGAGCGCTCCG 420

QY 141 AlaArgProProProProProProAlaSerValSerProGlnAlaGluProValTrpThr 160
DB 421 GCCCGGCTCCCTCTCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480

QY 161 ProProAlaProAlaProAlaProProSerThrProAlaAlaProLysArgArgGly 180
DB 481 CCGCCAGCCCGGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 540

QY 181 SerSerGlySerValValValAspLeuLeuTyrTrpArgAspLysLysThrGlyVal 200
DB 541 TCTCTGGGCTCAGTGTGTGTGACCTCTCTGAGGAGACATTAGGAGACTGGAGTG 600

QY 201 ValPheGlyAlaSerLeuPheLeuLeuLeuSerLeuThrValPheSerIleValSerVal 220
DB 601 GTGTTTGTGGCCAGCTATTCTCTGCTCTTTCATTGACAGTATTGACGATTGTGACGTA 660

QY 221 ThrAlaTyrIleAlaLeuAlaLeuLeuSerValThrIleSerProArgIleTyrLysGly 240
DB 661 ACAGCCTACATTCCCTTGGCCCTGCTCTCTGTGACCATCAGCTTTTAGGATATACAAGGT 720

QY 241 ValIleGlnAlaIleGlnLysSerAspGluGlyHisProPheArgAlaTyrLeuGluSer 260
DB 721 GTGATCCAGCTATCCAGAAATCAGATGAAGGCCACCCATTGAGGCAATATCTGGAATCT 780

QY 261 GluValAlaIleSerGluGluLeuValGlnLysTyrSerAsnSerAlaLeuGlyHisVal 280
DB 921 GTGATCCAGCTATCCAGAAATCAGATGAAGGCCACCCATTGAGGCAATATCTGGAATCT 980

; Sequence 5, Application US/09765205
; Patent No. US20020034800A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Li
; TITLE OF INVENTION: BONE MARROW SECRETED PROTEINS AND POLYNUCLEOTIDES
; FILE REFERENCE: 1459.004/200130.449
; CURRENT APPLICATION NUMBER: US/09/765,205
; CURRENT FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US/09/212,440
; PRIOR FILING DATE: 1998-12-16
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 1610
; TYPE: DNA
; ORGANISM: human
US-09-765-205-5

Alignment Scores:
Pred. No.: 1,01e-158 Length: 1610
Score: 1879.00 Matches: 370
Percent Similarity: 99.20% Conservatives: 0
Best Local Similarity: 99.20% Mismatches: 3
Query Match: 98.38% Indels: 0
DB: 9 Gaps: 0

US-09-544-776-2 (1-373) x US-09-765-205-5 (1-1610)

QY 1 MetGluAspLeuAspGlnSerProLeuValSerSerSerSerProProArgProGln 20
DB 132 ATGGAAGACTGGACCAAGTCTCTCTGGTCTCGTCTCGGACAGCCACCCCGGCGCAG 191

QY 21 ProAlaPheLysTyrGlnPheValArgGluProGluAspGluGluGluGluGlu 40
DB 192 CCGCGCTTCAAGTACCAGTTCTGTGGAGGAGCCCGGAGGAGGAGGAGGAGGAG 251

QY 41 GluGluGluAspGluAspGluAspGluGluGluGluLeuGluValLeuGluArgLysProAla 60
DB 252 GAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 311

QY 61 AlaGlyLeuSerAlaAlaProValProThrAlaProAlaAlaGlyAlaProLeuMetAsp 80
DB 312 GCCCGGCTCTCGCGCGCCCGAGTGGCCACCGCCCTCGCGCGCGCGCGCGCGCGCG 371

QY 81 PheGlyAsnAspPheValProProAlaProAlaProArgGlyPheLeuProAlaAlaProVal 100
DB 372 TTCGGAATGACTTCTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 431

QY 101 AlaProGluArgGlnProSerTrpAspProSerProValSerSerThrValProAlaPro 120
```

```
Db 432 GCGCCGAGCGGAGCGCGTCTTGGGACCGAGCGCGTCTGTCGACCGCGCGGCCA 491
QY 121 SerPheLeuSerAlaAlaValSerProSerLysLeuProGluAspGluProPro 140
Db 492 TCCCGCGTGTCTGCTGCGCGAGTCTCGCCCTCCAGCTCCCTGAGGACGAGCGCTCCG 551
QY 141 AlaArgProProProProProAlaSerValSerProGlnAlaGluProValTrpThr 160
Db 552 GCGCGGCTCCCGCTCTCCCGCGCGCGAGCGTGGAGCGCGCGCGCGTGGAGC 611
QY 161 ProProAlaProAlaProAlaProProSerThrProAlaAlaProLysArgArgGly 180
Db 612 CCGCCAGCGCGCGTCCCGCGCGCGCGCGCTCCAGCGCGCGCGCGCGCGCGCG 671
QY 181 SerSerGlySerValValValAspLeuLeuLeuLeuLeuLeuLeuLeuLeuVal 200
Db 672 TCTCGGCTCAGTGGTGTGTGACCTCTCTGTACTGGAGACATTAAGAAGACTGGAGTG 731
QY 201 ValPheGlyAlaSerLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuVal 220
Db 732 GTGTTGGTCCAGCTATTCTCTGCTGCTTTCATTGACAGTATTGAGCGTA 791
QY 221 ThrAlaTyrIleAlaLeuAlaLeuLeuSerValThrIleSerProArgIleTyrLysGly 240
Db 792 ACAGCCTACATTGCTTGGCGCTGCTCTCTGTGACCATCAGCTTTAGGATATACAAGGGT 851
QY 241 ValIleGlnAlaIleGlnLysSerAspGluGlyHisProPheArgAlaTyrLeuGluSer 260
Db 852 GTATCCAGCTATCCAGAAATCAGATGAAGGCCACCATTCAGGCGCATATCTGGAACTCT 911
QY 261 GluValAlaIleSerGluGluLeuValGlnLysTyrSerAsnSerAlaLeuGlyHisVal 280
Db 912 GAAGTTGCTATATCTGAGGAGTTGTTTTCAGAGTACAGTAATCTCTGCTTGTGTCATGTC 971
QY 281 AsnCysThrIleLysGluLeuArgArgLeuPheLeuValAspAspLeuValAspSerLeu 300
Db 972 AACTGCACGATAAAGGAACCTCAGCGCGCTCTCTCTAGTGTGATTTAGTTGATTTCTCTG 1031
QY 301 LysPheAlaValLeuMetTrpValPheThrTyrValGlyAlaLeuPheAsnGlyLeuThr 320
Db 1032 AAGTTTCAGTGTGTGATGGGTATTTTACTATGTTGGTCCCTGTTTAAATGTCGTGACA 1091
QY 321 LeuLeuIleLeuAlaLeuIleSerLeuPheSerValProValIleTyrGluArgHisGln 340
Db 1092 CTACTGATTTTGGCTCTCATTTTCACTTCTCAGTGTCTCTGTTTATTAAGACGCGATCAG 1151
QY 341 AlaGlnIleAspHisTyrLeuGlyLeuAlaAsnLysAsnValLysAspAlaMetAlaLys 360
Db 1152 GCACAGATAGATCATTTATCTAGGACTTGCAAATAAGAAATGTTAAAGATGCTATGGCTAA 1211
QY 361 IleGlnAlaLysIleProGlyLeuLysArgLysAlaGlu 373
Db 1212 ATCCAAAGCAAAATCCCTGGATTGAAGCGCAAGCTGAA 1250
```

## RESULT 3

```
US-10-347-669-5
; Sequence 5, Application US/10347669
; Publication No. US20050084850A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Li
; TITLE OF INVENTION: BONE MARROW SECRETED PROTEINS AND POLYNUCLEOTIDES
; FILE REFERENCE: 1458.004/200130.449
; CURRENT APPLICATION NUMBER: US/10/347,669
; CURRENT FILING DATE: 2003-01-16
; PRIOR APPLICATION NUMBER: US/09/212,440
; PRIOR FILING DATE: 1998-12-16
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 1610
; TYPE: DNA
; ORGANISM: human
```

```
US-10-347-669-5
Alignment Scores:
Pred. No.: 1,01e-158 Length: 1610
Score: 1879.00 Matches: 370
Percent Similarity: 99.20% Conservative: 0
Best Local Similarity: 99.20% Mismatches: 3
Query Match: 98.38% Indels: 0
DB: 21 Gaps: 0

US-09-544-776-2 (1-373) x US-10-347-669-5 (1-1610)
QY 1 MetGluAspLeuAspGlnSerProLeuValSerSerAspSerProProArgProGln 20
Db 132 ATGGAAGACCTGGACCACTCTCTCTGCTCTCTCTCGACAGCCACCCGCGCGCGCAG 191
QY 21 ProAlaPheIstYrGlnPheValArgGluProGluAspGluGluGluGluGlu 40
Db 192 CCGCGCTTCAAGTACCAGTTCTGTGAGGAGCGCGCGAGGAGGAGGAGGAG 251
QY 41 GluGluGluAspGluAspGluAspLeuGluGluGluValLeuGluArgLysProAla 60
Db 252 GAAGAGGAGGAGCGAGGACGAGACCTGGAGGAGCTGGAGGTGCTGGAGAGGAGCGCGCC 311
QY 61 AlaGlyLeuSerAlaAlaProValProThrAlaProAlaAlaGlyAlaProLeuMetAsp 80
Db 312 GCGCGCTGTCCGCGCGCGCGAGTGCACCGCCCTGCGCGCGCGCGCGCTGATGGAC 371
QY 81 PheGlyAsnAspPheValProProAlaProAlaProArgGlyPheLeuProAlaProVal 100
Db 372 TTCGAAATATGACTTCTGCGCGCGCGCGCGCGCGCGCGCGCGCTGCCCGCGCTCCCCCGTC 431
QY 101 AlaProGluArgGlnProSerTrpAspProSerProValSerSerThrValProAlaPro 120
Db 432 GCGCGCGAGCGCGCGCGCTCTTGGGACCGCGCGCGCTGCTGCTGCGCGCGCGCGCGCA 491
QY 121 SerPheLeuSerAlaAlaValSerProSerLysLeuProGluAspGluProPro 140
Db 492 TCCCGCTGTCTGCTCGCGCGCTCTCCAGCTCTCCAGCTCTCCAGGACGAGCGCGCTCCG 551
QY 141 AlaArgProProProProProProAlaSerValSerProGlnAlaGluProValTrpThr 160
Db 552 GCGCGCGCTTCCCGCTCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGTGTGACC 611
QY 161 ProProAlaProAlaProAlaProProSerThrProAlaAlaProLysArgArgGly 180
Db 612 CCGCGAGCGCGCGCTCCCGCGCGCGCGCTCCAGCGCGCGCGCGCGCGCGCGCGCGCG 671
QY 181 SerSerGlySerValValValAspLeuLeuLeuTyrTrpArgAspIleLysThrGlyVal 200
Db 672 TCTCGGCTCAGTGGTGTGTGACCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 731
QY 201 ValPheGlyAlaSerLeuPheLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuVal 220
Db 732 GTGTTTGGTGGCGAGCTATTCTCTGCTGCTTTCATTGACAGTATTGACATTTGAGCGTA 791
QY 221 ThrAlaTyrIleAlaLeuAlaLeuLeuSerValThrIleSerProArgIleTyrLysGly 240
Db 792 ACAGCCTACATTGCTTGGCGCTGCTCTCTGTGACCATCAGCTTTAGGATATACAAGGGT 851
QY 241 ValIleGlnAlaIleGlnLysSerAspGluGlyHisProPheArgAlaTyrLeuGluSer 260
Db 852 GTGATCCAGCTATCCAGAAATCAGATGAAGGCCACCATTCAGGCGCATATCTGGAACTCT 911
QY 261 GluValAlaIleSerGluGluLeuValGlnLysTyrSerAsnSerAlaLeuGlyHisVal 280
Db 912 GAAGTTGCTATATCTGAGGAGTTGTTTTCAGAGTACAGTAATCTCTGCTTGTGTCATGTC 971
QY 281 AsnCysThrIleLysGluLeuArgArgLeuPheLeuValAspAspLeuValAspSerLeu 300
Db 972 AACTGCACGATAAAGGAACCTCAGCGCGCTCTCTCTAGTGTGATTTAGTTGATTTCTCTG 1031
QY 301 LysPheAlaValLeuMetTrpValPheThrTyrValGlyAlaLeuPheAsnGlyLeuThr 320
```

Db	1032	AAAGTTTCAGTGTGGTATTTACCTATATTGGTGCCTGTTTAATGGTCTGACA	10931
Qy	321	LeuLeuIleLeuAlaLeuIleSerLeuPheSerValProValIleTyrGluAtgHisGln	340
Db	1092	CTACTGATTTGGCTCTCAATTTCACTCTTCAAGTGTCCTGTATTATGAACGGCATCAG	1151
Qy	341	AlaGlnIleAspHisTyrLeuGlyLeuAlaAsnLysAsnValLysAspAlaMetAlaLys	360
Db	1152	GCACAGATAGATCATTTATCTAGGACTTGCAAAATAAGAAATGTTAAAGATGCTATGGCTAAA	1211
Qy	361	IleGlnAlaLysIleProGlyLeuLysArgLysAlaGlu	373
Db	1212	ATCCAAGCAAAAATCCCTGGATTGAAGCGCAAGCTGAA	1250

RESULT 4

US-10-956-157-4530

; Sequence 4530, Application US/10956157

; Publication No. US20050118625A1

; GENERAL INFORMATION:

; APPLICANT: Wyeth

; APPLICANT: Mounts, William

; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH

; TITLE OF INVENTION: HUMAN OSTEOARTHRITIS AND HUMAN PROTEASES

; FILE REFERENCE: 031896-043000 (AM 101081)

; CURRENT APPLICATION NUMBER: US/10/956,157

; CURRENT FILING DATE: 2004-10-04

; NUMBER OF SEQ ID NOS: 319805

Query Match:	98.38%	Indels:	0
DB:	21	Gaps:	0
US-09-544-776-2 (1-373) x US-10-956-157-4531 (1-2226)			
QY	1	MetGluAspLeuAspGlnSerProLeuValSerSerSerAspSerProProArgProGln	20
DB	142	ATGGAAGACCTGGACAGCTCTCTGCTCGTCTCGTCTCGACAGCCACCCCGCGCGCAG	201
QY	21	ProAlaPheLysTyrGlnPheValArgGluProGluAspGluGluGluGluGlu	40
DB	202	CCCGCGTTCAGTACCAAGTTCGTGAGGAGCCCGAGGACGAGGAGGAGGAGGAG	261
QY	41	GluGluGluAspGluAspGluAspLeuGluGluGluGluGluGluGluGluGlu	60
DB	262	GAAGAGGAGGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	321
QY	61	AlaGlyLeuSerAlaAlaProValProThrAlaProAlaAlaGlyAlaProLeuMetAsp	80
DB	322	GCCGGGCTGTCCGGCGCCCGAGTGCCTCCAGCTCCCGCGCGCGCCCTGATGGAC	381
QY	81	PheGlyAsnAspPheValProAlaProArgGlyPheLeuProAlaProVal	100
DB	382	TTCCGAAATGACTTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGTC	441
QY	101	AlaProGluArgGlnProSerProSerProSerProSerProSerProSerProSerPro	120
DB	442	GCCCCGGAGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC	501
QY	121	SerPheLeuSerAlaAlaAlaValSerProSerLysLeuProGluAspGluProPro	140
DB	502	TCCCGCGCTGTCTGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC	561
QY	141	AlaArgProProProProProProProProProProProProProProProProPro	160
DB	562	GCCCGCGCTCCCGCT	621
US-09-544-776-2 (1-373) x US-10-060-036-54 (1-2235)			
QY	1	MetGluAspLeuAspGlnSerProLeuValSerSerSerAspSerProProArgProGln	20
DB	142	ATGGAAGACCTGGACAGCTCTCTGCTCGTCTCGTCTCGACAGCCACCCCGCGCGCAG	201
QY	21	ProAlaPheLysTyrGlnPheValArgGluProGluAspGluGluGluGluGlu	40
DB	202	CCCGCGTTCAGTACCAAGTTCGTGAGGAGCCCGAGGAGGAGGAGGAGGAGGAG	261
QY	41	GluGluGluAspGluAspGluAspLeuGluGluGluGluGluGluGluGluGlu	60
DB	262	GAAGAGGAGGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	321
QY	61	AlaGlyLeuSerAlaAlaProValProThrAlaProAlaAlaGlyAlaProLeuMetAsp	80
DB	322	GCCGGGCTGTCCGGCGCCCGAGTGCCTCCAGCTCCCGCGCGCGCGCGCGCGCGC	381
QY	81	PheGlyAsnAspPheValProProAlaProArgGlyPheLeuProAlaProProVal	100
DB	382	TTCCGAAATGACTTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGTC	441
QY	101	AlaProGluArgGlnProSerProSerProSerProSerProSerProSerProSerPro	120
DB	442	GCCCCGGAGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC	501
QY	121	SerPheLeuSerAlaAlaAlaValSerProSerLysLeuProGluAspGluProPro	140
DB	502	TCCCGCGCTGTCTGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC	561
QY	141	AlaArgProProProProProProProProProProProProProProProProPro	160
DB	562	GCCCGCGCTCCCGCT	621
US-10-060-036-54			
; Sequence 54, Application US/10060036			
; Publication No. US20030073144A1			
; GENERAL INFORMATION:			
; APPLICANT: Benson, Darin R.			
; APPLICANT: Kalos, Michael D.			
; APPLICANT: Lodes, Michael J.			
; APPLICANT: Persing, David H.			
; APPLICANT: Hepler, William T.			
; APPLICANT: Jiang, Yuqiu			
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY			
; FILE OF INVENTION: AND DIAGNOSIS OF PANCREATIC CANCER			
; CURRENT APPLICATION NUMBER: US/10/060,036			
; CURRENT FILING DATE: 2002-01-30			
; NUMBER OF SEQ ID NOS: 4560			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 54			
; LENGTH: 2235			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
US-10-060-036-54			
Alignment Scores:			
Pred. No.:	1,56e-158	Length:	2235
Score:	1879.00	Matches:	370
Percent Similarity:	99.20%	Conservative:	0
Best Local Similarity:	99.20%	Mismatches:	3
Query Match:	98.38%	Indels:	0
DB:	14	Gaps:	0

```
QY 161 ProProAlaProAlaProAlaProAlaProSerThrProAlaAlaProLysArgArgGly 180
Db 622 CGCCAGCCCCGGCTCCCGCGGCCCTCCACCCCGCGCGCCCAAGCGCAGGGGC 681
QY 181 SerSerGlySerValValValAspLeuLeuLeuTyrTrpArgAspIleLysLysThrGlyVal 200
Db 682 TCCTCGGGCTCAGTGGTGTGTGACCTCTCTGACTGGAGAGACATTAAAGAAGACTGGAGTG 741
QY 201 ValPheGlyAlaSerLeuPheLeuLeuLeuSerLeuThrValPheSerIleValSerVal 220
Db 742 GTGTTTGGTGGCCAGCCATTCTCTGCTGCTTTTCATTGACAGTATTCAGCAATTGTGAGCGTA 801
QY 221 ThrAlaTyrIleAlaLeuAlaLeuLeuSerValThrIleSerProArgIleTyrLysGly 240
Db 802 ACAGCTACATTGCTTGGCCCTGCTCTCTGTGACCATCAGCTTTAGGATATACAAAGGT 861
QY 241 ValIleGlnAlaIleGlnLysSerAspGluGlyHisProPheArgAlaTyrLeuGluSer 260
Db 862 GTGATCCAAAGCTATCCAGAAATCAGATGAAGGCCACCCATTTCAGGGCATATCTGGAATCT 921
QY 261 GluValAlaIleSerGluGluLeuValGlnLysTyrSerAsnSerAlaLeuGlyHisVal 280
Db 922 GAAGTTGCTATATCTGAGGAGTTGGTTCAAGAGTACAGTAATTCCTCTTGGTCATGTG 981
QY 281 AsnCysThrIleLysGluLeuArgArgLeuPheLeuValAspAspLeuValAspSerLeu 300
Db 982 AACTGACAGATTAAGGAATCAGCGCCCTCTCTTAGTTGATGATTTAGTTGATCTCTG 1041
QY 301 LysPheAlaValLeuMetTrpValPheThrTyrValGlyAlaLeuPheAsnGlyLeuThr 320
Db 1042 AAGTTTGCAAGTGTGATGGGTATTTACCTATGTTGGTGCCTGTTTAATGGTCTGACA 1101
QY 321 LeuLeuIleLeuAlaLeuIleSerLeuPheSerValProValIleTyrGluArgHisGln 340
Db 1102 CTACTGATTTGGCTCTCATTTTCACTCTTCACTGTTTCTCTGTTTATTTATGAACGGCATCAG 1161
QY 341 AlaGlnIleAspHisTyrLeuGlyLeuAlaAsnLysAsnValLysAspAlaMetAlaLys 360
Db 1162 GCACGATAGATCATTAATCAGACTTTCAGAAATAGAAATGTTAAGATGCTATGGCTAAA 1221
QY 361 IleGlnAlaLysIleProGlyLeuLysArgLysAlaGlu 373
Db 1222 ATCCAAGCAAAATCCCTGGATTGAAGCGCAAGCTGAA 1260
```

## RESULT 7

```
US-10-466-258-3
; Sequence 3, Application US/10466258
; Publication No. US20040132096A1
; GENERAL INFORMATION:
; APPLICANT: GLAXO GROUP LIMITED
; TITLE OF INVENTION: ASSAY
; FILE REFERENCE: P80966 GCW
; CURRENT APPLICATION NUMBER: US/10/466,258
; CURRENT FILING DATE: 2003-07-15
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 2052
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (67)..(1188)
US-10-466-258-3
```

## Alignment Scores:

Pred. No.:	7,31e-158	Length:	2052
Score:	1871.00	Matches:	368
Percent Similarity:	98.93%	Conservatives:	1
Best Local Similarity:	98.66%	Mismatches:	4
Query Match:	97.96%	Indels:	0
DB:	19	Gaps:	0

```
US-09-544-776-2 (1-373) x US-10-466-258-3 (1-2052)
QY 1 MetGluAspLeuAspGlnSerProLeuValSerSerAspSerProProArgProGln 20
Db 67 ATGGAAGACCTGGACAGTCTCTCTGGTCTCTCGACAGCCACCCCGCGCCAG 126
QY 21 ProAlaPheLysTyrGlnPheValArgGluProGluAspGluGluGluGluGlu 40
Db 127 CCCGGTTCAGTACAGTTCTGTGAGGAGCCCGAGGACGAGGAGGAGGAGGAG 186
QY 41 GluGluGluAspGluAspGluAspLeuGluGluLeuGluValLeuGluArgLysProAla 60
Db 187 GAAGAGGAGGACGAGGACGAGACACCTGGAGGAGCTGGAGTCTGTGAGAGGAGCCCGC 246
QY 61 AlaGlyLeuSerAlaAlaProValProThrAlaProAlaAlaGlyAlaProLeuMetAsp 80
Db 247 GCCGGGCTGTCCGGCGCCCACTGAGTGCACCGCCCTGCGCGCGCGCGCCCTGATGGAC 306
QY 81 PheGlyAsnAspPheValProProAlaProArgGlyPheLeuProAlaAlaProVal 100
Db 307 TTCGGAATGATCTTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTCCCCCGCTC 366
QY 101 AlaProGluArgGlnProSerTrpAspProSerProValSerSerThrValProAlaPro 120
Db 367 GCCCGGAGCGGCACCGCTGTTGGGACCCGAGCCCGGTGCTCGACCGTGCCTCGGCCA 426
QY 121 SerPheLeuSerAlaAlaValSerProSerLysLeuProGluAspAspGluProPro 140
Db 427 TCCCGCTGTCTGCTGCGCGAGTCTCGCCCTCAAGCTCCCTCAGGACGACGAGCTCGG 486
QY 141 AlaArgProProProProProAlaSerValSerProGlnAlaGluProValTrpThr 160
Db 487 GCCCGCTTCCCTCTCTCCCGCGCGAGCGTGAGCCCCCGAGGAGCCCGTGTGGACC 546
QY 161 ProProAlaProAlaProProSerThrProAlaAlaProLysArgArgGly 180
Db 547 CGCCAGCCCCGGCTCCCGCGCGCCCTCCACCCCGCGCGCCCAAGCGCAGGGGC 606
QY 181 SerSerGlySerValValValAspLeuLeuTyrTrpArgAspIleLysLysThrGlyVal 200
Db 607 TCCTCGGGCTCAGTGGTGTGTGACCTCTCTGACTCGAGAGACATTAAAGAAGACTGGAGTG 666
QY 201 ValPheGlyAlaSerLeuPheLeuLeuLeuSerLeuThrValPheSerIleValSerVal 220
Db 667 GTGTTTGGTGGCAGGCTTATCTCTGCTGCTTTCATTGACAGTATTCAGCATTTGTGAGCGTA 726
QY 221 ThrAlaTyrIleAlaLeuAlaLeuLeuSerValThrIleSerProArgIleTyrLysGly 240
Db 727 ACAGCTACATTGCTTGGCCCTGCTCTCTGTGACCATCAGCTTTTAGGATATACAAAGGT 786
QY 241 ValIleGlnAlaIleGlnLysSerAspGluGlyHisProPheArgAlaTyrLeuGluSer 260
Db 787 GTGATCCAAGCTATATCCAGAAATCAGATGAAGGCCACCCATTTCAGGGCATATCTGGAATCT 846
QY 261 GluValAlaIleSerGluGluLeuValGlnLysTyrSerAsnSerAlaLeuGlyHisVal 280
Db 847 GAAGTTGCTATATCTGAGAGTGTGTTTCAAGATACAGTAATTCCTCTTGGTCTATGTG 906
QY 281 AsnCysThrIleLysGluLeuArgArgLeuPheLeuValAspAspLeuValAspSerLeu 300
Db 907 AACTGCACGATAAAGAACTCAGCGCGCTCTTCTTAGTTGATGATTTAGTTCATTCTCTG 966
QY 301 LysPheAlaValLeuMetTrpValPheThrTyrValGlyAlaLeuPheAsnGlyLeuThr 320
Db 967 AAGTTTGCAGTGTTCATGTGGGTATTTACCTATGTTGGTGCCTTGTAAATGGTCTGACA 1026
QY 321 LeuLeuIleLeuAlaLeuIleSerLeuPheSerValProValIleTyrGluArgHisGln 340
Db 1027 CTACTGATTTGGCTCTCATTTCACTCTTCACTGTTCTCTGTTTATTTATGAACGGCATCAG 1086
QY 341 AlaGlnIleAspHisTyrLeuGlyLeuAlaAsnLysAsnValLysAspAlaMetAlaLys 360
Db 1086
```





[illegible]





Db 2581 GATTTCATCTCCAATTGAATATAGATGAGTTCCTACATGATCAGTTCATAAACTGAT 2640  
QY 185 ----- 185  
Db 2641 TCATTTTCTAAATTAGCCAGGAATATAGTACCTAGATATCCCAAAAGTGAAAT 2700  
QY 185 ----- 185  
Db 2701 GCTAATGCCCGGATGGAGCTGGTCAATTCGCTTGACAGAAATGGCCCATGACCTTTCT 2760  
QY 185 ----- 185  
Db 2761 TTGAAGAACATACAACCCAAAGTTGAAGAGAAAAATCAGTTTCTCAGATGACTTTTCTAAA 2820  
QY 185 ----- 185  
Db 2821 AATGGGTCTGCTACATCAAAAGTGCTTTATTGTCCTCCAGATGTTTCTGCTTTGGCCACT 2880  
QY 185 ----- 185  
Db 2881 CAAGCAGATAGAGAGCATAGTTAAACCAAAAGTTCTTGTAAGAAAGCTGAGAAAAA 2940  
QY 185 ----- 185  
Db 2941 CTTCTCTCCGATACAGAAAAAGAGGACAGATCACCATCTGCTATATTTTCAGCAGAGCTG 3000  
QY 186 -----ValValAspLeuLeuTyrTrpArgAspIleIleIysThrGlyValVal 201  
Db 3001 AGTAAACTTCAGTTGTTGACCTCTCTGTACTGGAGAGACATTAAAGAAGACTGGAGTGGT 3060  
QY 202 PheGlyValSerLeuPheLeuLeuSerLeuThrValPheSerIleValSerValThr 221  
Db 3061 TTTGGTGCAGCCTATCTCTGCTTTTCATTGACAGATATTGACCATTTGTGAGCGTAACA 3120  
QY 222 AlaTyrIleAlaLeuAlaLeuSerValThrIleSerProArgIleTyrIysGlyVal 241  
Db 3121 GCCTACATTTGCCCTTGCTCTCTGTGACCATCAGCTTTTAGGATATACAAAGGTGTG 3180  
QY 242 IleGlnAlaIleGlnIysSerAspGluGlyHisProPheArgAlaTyrIleuGluSerGlu 261  
Db 3181 ATCCAAGCTATCCAGAAATCAGATGAAGGCCACCCCATTCAGGGCATATCTGGAATCTGAA 3240  
QY 262 ValAlaIleSerGluGluLeuValGlnIysTyrSerAsnSerAlaLeuGlyHisValAsn 281  
Db 3241 GTTGCTATATCTGAGAGTTGGTTCAAGAAGTACAGTAATCTGCTTGTGTCATGTGAAC 3300  
QY 282 CysThrIleIysGluLeuArgLeuPheLeuValAspLeuValAspSerLeuIys 301  
Db 3301 TGCACGATAAAGAACTCAGGGGCCCTCTTCTTAGTTGATGATTTAGTTGATTTCTCTGAAG 3360  
QY 302 PheAlaValLeuMetTrpValPheThrTyrValGlyAlaLeuPheAsnGlyLeuThrIleu 321  
Db 3361 TTTGCAGTGTGATGGGTATTTACCTATGTTGGTGCCCTTGTTAATGGTCTGACACTA 3420  
QY 322 LeuIleLeuAlaLeuIleSerLeuPheSerValProValIleTyrGluArgHisGlnAla 341  
Db 3421 CTGATTTTGGCTCTCATTTTCATCTTCAGTGTTCCTGTTATTATGAAGCGATCAGGGC 3480  
QY 342 GlnIleAspHisTyrLeuGlyLeuAlaAsnIysAsnValIysAspAlaMetAlaIysIle 361  
Db 3481 CAGATAGATCATTTATCTAGACTTGCAAAATAAGAATGTTAAAGATGCTATGCTGCTAAAATC 3540  
QY 362 GlnAlaIleProGlyLeuIysArgIysAlaGlu 373  
Db 3541 CAAGCAAAAAATCCCTGGATTGAAGCCAAAGCTGAA 3576

## RESULT 11

US-09-893-348-22

; Sequence 22, Application US/09893348

; Patent No. US20020072493A1

; GENERAL INFORMATION:

; APPLICANT: EISENBACH-SCHWARTZ, Michal

; APPLICANT: COHEN, Irun R.

; APPLICANT: BESERMAN, Pierre  
; APPLICANT: MOSONEGO, Alon  
; APPLICANT: MOALEM, Gila  
; TITLE OF INVENTION: ACTIVATED T-CELLS, NERVOUS SYSTEM-SPECIFIC ANTIGENS AND THEIR USE  
; FILE REFERENCE: EIS-SCHWARTZ-2A  
; CURRENT APPLICATION NUMBER: US/09/893,348  
; CURRENT FILING DATE: 2001-06-28  
; PRIOR APPLICATION NUMBER: US 09/314,161  
; PRIOR FILING DATE: 1999-05-19  
; PRIOR APPLICATION NUMBER: US 09/218,277  
; PRIOR FILING DATE: 1998-12-22  
; PRIOR APPLICATION NUMBER: PCT/US98/14715  
; PRIOR FILING DATE: 1998-07-21  
; PRIOR APPLICATION NUMBER: IL 124500  
; PRIOR FILING DATE: 1998-05-19  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 22  
; LENGTH: 3579  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(3579)  
; OTHER INFORMATION:  
US-09-893-348-22

## Alignment Scores:

Pred. No.: 1.38e-120 Length: 3579  
Score: 1459.50 Matches: 370  
Percent Similarity: 31.04% Conservative: 0  
Best Local Similarity: 31.04% Mismatches: 3  
Query Match: 76.41% Indels: 819  
DB: 9 Gaps: 1

US-09-544-776-2 (1-373) x US-09-893-348-22 (1-3579)

QY 1 MetGluAspLeuAspGlnSerProLeuValSerSerAspSerProProArgProGln 20  
Db 1 ATGGAAGACCTGGACAGTCTCTCTGTGCTCTGTCTCGACAGCCACCCCGGCGCAG 60  
QY 21 ProAlaPheLysTyrGlnPheValArgGluProGluAspGluGluGluGluGlu 40  
Db 61 CCCGGTTCAAGTACCAGTTCTGTAGGGAGCCCGAGGACGAGGAGGAGGAG 120  
QY 41 GluGluGluAspGluAspGluAspLeuGluGluLeuGluValLeuGluArgLysProAla 60  
Db 121 GAAGAGGAGGACGAGAGCGAAGACCTGGAGGAGCTGGAGGTCTGGAGAGGAAGCCCGCC 180  
QY 61 AlaGlyLeuSerAlaAlaProValProThrAlaProAlaAlaGlyAlaProLeuMetAsp 80  
Db 181 GCCGGGTGTCTCGCGGCCCGCCAGTCCACCGCCCTCGCGCGCGCGCCCTGTATGAC 240  
QY 81 PheGlyAsnAspPheValProProAlaProArgGlyPheLeuProAlaAlaProVal 100  
Db 241 TTCGGAATGACTTCTGTGCGCGCGCGCGCGCGCGCGCGCGCGCTCCCCCGCTC 300  
QY 101 AlaProGluArgGlnProSerTrpAspProSerProValSerSerThrValProAlaPro 120  
Db 301 GCCCGGAGCGGACACCGCTTCTGGACCCGAGCCCGGTGTCTGTCACCGTGTCCCGGCCA 360  
QY 121 SerPheLeuSerAlaAlaAlaValSerProSerLysLeuProGluAspAspGluProPro 140  
Db 361 TCCCCGTGTCTGTCTGCGCGAGTCTCGCCCTCCAGCTCCCTGAGGACGACGAGCTCG 420  
QY 141 AlaArgProProProProProAlaSerValSerProGlnAlaGluProValTrpThr 160  
Db 421 GCCCGGCTTCCCCCTCTCCCCCGCCAGCGGTGAGCCCGCCAGGACAGCCCGTGTGAC 480  
QY 161 ProProAlaProAlaProAlaProProSerThrProAlaAlaProLysArgGly 180  
Db 481 CGCCAGCCCGGGTCTCCCGCGCGCCCTCTCCACCCCGCGCGCGCCCAAGCGAGGGGC 540

181 SerSertGlySerVal----- 185  
| | | | | | | | | |  
541 TCCTCGGGCTCAGTGGATGAGACCCCTTTTCTCTCTGCTGCATCTGAGCCTGTGATA 600  
185 ----- 185  
601 CGTCTCTCTCAGAAATATGGACTTGAAGAGCAGCCAGGTAACTATTTCCGGTGGT 660  
185 ----- 185  
661 CAAGAGGATTTCCCATCTGCTGCTTGAACCTGCTGCTTCTCTCTCTCTCTCTCTCT 720  
185 ----- 185  
721 CTCTCAGCCGCTTCTTTCAAGAACAATGAAATACCTTGGTAAATTTGTCAACAGTATTACCC 780  
185 ----- 185  
781 ACTGAAGGAACACTTCAAGAAATGTCAAGTAAAGCTTTTAAAGAGGTCTCAGAGAAGGCA 840  
185 ----- 185  
841 AAACTCTACTCATAGATAGAGATTTAAACAGAGTTTTCAGAAATTAGAAATACTCAGAAATG 900  
185 ----- 185  
901 GGATCATCGTTTCACTGCTCTCCAAAAGCAGAACTCTGCCGTAATAGTAGCAAAATCCTAGG 960  
185 ----- 185  
961 GAAGAAATAATCGTGAATAATAAGATGAAGAAGAGAGTGTAGTTAGTAATAACATCCTT 1020  
185 ----- 185  
1021 CATAATCAACAGAGTTTACCTACAGCTCTTACTAAATGTTTAAAGAGGATGAAGTTGTG 1080  
185 ----- 185  
1081 TCTTCAGAAAAAGCAAAAGACAGTTTTTAATGAAAAGAGAGTTTGCAGTGGAAAGCTCCTATG 1140  
185 ----- 185  
1141 AGGAGGAATATGCAGACTTCAAACCATTTTGAAGAGTATGGGAAGTGAAGATAGTAAG 1200  
185 ----- 185  
1201 GAAGATAGTATATGTTGGCTGCTGGAGGTAAATTCGAGAGCAACTTGGAAAGTAAAGTG 1260  
185 ----- 185  
1261 GATAAAAAATGTTTTCAGATAGCTTTCAGCAAACTAAATCAGCAAAAAGATAGTGAGAGT 1320  
185 ----- 185  
1321 AGTAATGATGATACTTCTTTCCAGTACGCCAGAGGTATAAAGGATCGTCCAGGAGCA 1380  
185 ----- 185  
1381 TATATCACATGTGCTCCCTTTAACCCAGCAGCAACTGAGAGCATTTGCAACAAACATTTTT 1440  
185 ----- 185  
1441 CCTTTGTTAGGATCCTACTTCAGAAATAAGACCGATGAAAAAATAAGAAAAAG 1500  
185 ----- 185  
1501 AAGSCCAATAGTAACAGAGAAGAACTACTAGCACCAAAACATCAAAACCCTTTTCTTTGTA 1560  
185 ----- 185  
1561 GCAGCACAGGATTTCTGAGACAGATTATGTCAACACAGATAATTTTAACAAAGGTGACTGAG 1620  
185 ----- 185

1621 GAAGTCGTGGCAAAACATGCTGAAGCCCTGACTCCAGATTTAGTACAGGAAGCATGTGAA 1680  
185 ----- 185  
1681 AGTGAATTCGAATGAAGTTACTGTGTACAAAGATTGCTTATGAACAACAAAAATGGACTGGTT 1740  
185 ----- 185  
1741 CAAACATCAGAAAGTTATGCAAGAGTCACTCTATCTCTGAGCAGCAGACTTTTGCCCATCATTT 1800  
185 ----- 185  
1801 GAAGAGTCAAGAGTACTCTCTCACCAGTTTTCGCCGTGACATTTGTTATGGAAGCACCATTG 1860  
185 ----- 185  
1861 AATTCGACAGTTCCTTAGTGTCTGTGCTTCGGTGATACAGCCAGCTCATCACCATTAGAA 1920  
185 ----- 185  
1921 GCTTCTCTCAGTTAAATTTATGAAGCATATAAACAATGAGCCTGAAAAACCCCCACCATATGAA 1980  
185 ----- 185  
1981 GAGGCATGAGTGTATCACTAAAAAAGTATCAGGAATAAGGAAGAAAAATTAAGAGCCT 2040  
185 ----- 185  
2041 GAAATATTAAATGACAGCTCTTCAAGAAACAGAAAGCTCTTATATATATCTATTGTCATGTGAT 2100  
185 ----- 185  
2101 TTAATTTAAAGAAACAAAGCTTTTCTGCTGAACAGCTCCGATTTCTCTGATTATTACAGAA 2160  
185 ----- 185  
2161 ATGGCAAAAGTTGAACAGCAGCTGCTGATCACTTCTGAGCTAGTTGAAGATTCTCACCT 2220  
185 ----- 185  
2221 GATTCGAACCAAGTTGACTTATTATTAGTGATGATTCATTAATCTGACGTTCCACAAAAACAA 2280  
185 ----- 185  
2281 GATGAAACTGTGATGCTTGTGAAAGAAAGTCTCACTGAGACTTTCATTTTGAGTCAATGATA 2340  
185 ----- 185  
2341 GAATATGAAAAATAAGGAAAAAAGTCTGCTTTGCCACCTGAGGGAGGAAAGCCATATTTG 2400  
185 ----- 185  
2401 GAATCTTTTAAGCTCAGTTTAGATAACACAAAGATACCTGTTCCTGATGAAGTTTCA 2460  
185 ----- 185  
2461 ACATTGAGCAAAAAGGAGAAAAATTCCTTTGCAGATGGAGAGCTCAGTACTGCAGTTTAT 2520  
185 ----- 185  
2521 TCAATGATGACTTATTATTATTTCTAAGGAAGCAGACATAAGAGAACTGAAACGTTTCA 2580  
185 ----- 185  
2581 GATTCAATCTCCAATTGAAATTTATAGATGAGTTCCTTACATTGATGATGATCTTAAACTGAT 2640  
185 ----- 185  
2641 TCATTTTCTAAATTAGCCAGGGAATATCTGACCTAGAGATATCCCAACAAAAGTGAATTT 2700  
185 ----- 185

```
Db 2701 GCTAATGCCCGGATGGAGCTGGGTCAATTGCCCTTGCACAGAAATTGCCCATGACCTTTCT 2760
QY 185 ----- 185
Db 2761 TTGAAGAACATACAAACCAAGTTGAAGAGAAAAATCAGTTTCTCAGATGACTTTTCTAAA 2820
QY 185 ----- 185
Db 2821 AATGGGTCTGCTACATCAAAAGGTGCTCTATTGGCTCCAGATGTTTCTGTTGGCCACT 2880
QY 185 ----- 185
Db 2881 CAAGCAGAGATAGAGAGCATAGTTAAACCCAAAGTTCTTGTGAAGAAGCTGAGAAAAA 2940
QY 185 ----- 185
Db 2941 CTTCTTCCGATACAGAAAAAGAGGACAGATCACCATCTGCTATATTTTTCAGCAGAGCTG 3000
QY 186 -----ValValAspLeuLeuTyrTrpArgAspIleLysLysThrGlyValVal 201
Db 3001 AGTAAACCTTCAGTTGTTGACCTCTCTGTACTGGAGAGACATTAAGAAGACTGGAGTGGT 3060
QY 202 PheGlyAlaSerLeuPheLeuLeuSerLeuThrValPheSerIleValSerValThr 221
Db 3061 TTTGGTGCAGCCTATTCTCTGCTGCTTTCAATTGACAGTATTGAGCATTTGTGAGCGTAACA 3120
QY 222 AlaTyrIleAlaLeuAlaLeuSerValThrIleSerProArgIleTyrLysGlyVal 241
Db 3121 GCCTACATTGCCCTGGCCCTGCTCTGTGACCATCAGCTTTTAGGATATATACAAAGGTGTG 3180
QY 242 IleGlnAlaIleGlnLysSerAspGluGlyHisProPheArgAlaTyrLeuGluSerGlu 261
Db 3181 ATCCAGCTATCCAGAAATCAGATGAAGGCCACCCATTGAGGCATATCTGGAATCTGAA 3240
QY 262 ValAlaIleSerGluLeuValGlnLysTyrSerAsnSerAlaLeuGlyHisValAsn 281
Db 3241 GTTGCTATATCTGAGGAGTTGGTTCAAGAAGTACAGTAATTCGTCTTGCTCATGTGAAC 3300
QY 282 CysThrIleLysGluLeuArgIlePheLeuValAspIleValAspSerIleLys 301
Db 3301 TGACCATAAAGAACTCAGAGGCCCTCTCTTAGTTGATGATTAGTTGATTTCTTGAAG 3360
QY 302 PheAlaValLeuMetTrpValPheThrTyrValGlyAlaLeuPheAsnGlyLeuThrLeu 321
Db 3361 TTTGCACTGTGTATGGGTATTACTATGTGGTGCCTTGTTTAATGCTGACACTA 3420
QY 322 LeuIleAlaLeuLysSerLeuPheSerValProValIleTyrGluArgHisGlnAla 341
Db 3421 CTGATTTTGGCTCTCATTTTCACTCTTCAGTGTTCCTGTTATTATTATGAACGGCATCAGGCG 3480
QY 342 GlnIleAspHisTyrIleuGlyLeuAlaAsnLysAsnValLysAspAlaMetAlaLysIle 361
Db 3481 CAGATAGATCATTTATCTAGACTTGCATAATAAGAATGTTAAAGATGCTATGCTAAATATC 3540
QY 362 GlnAlaLysIleProGlyLeuLysArgLysAlaGlu 373
Db 3541 CAAGCAAAATCCCTGGATTGAAGCCAAAGCTGAA 3576
```

## RESULT 12

```
US-10-267-502-212
; Sequence 212, Application US/10267502
; Publication No. US20040071700A1
; GENERAL INFORMATION:
; APPLICANT: Kim, Jaeseob
; APPLICANT: Galant, Ron
; TITLE OF INVENTION: Obesity Linked Genes
; FILE REFERENCE: LSD-07416
; CURRENT APPLICATION NUMBER: US/10/267,502
; CURRENT FILING DATE: 2003-01-27
; NUMBER OF SEQ ID NOS: 439
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 212
; LENGTH: 3579
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-267-502-212
Alignment Scores:
Pred. No.: 1,38e-120 Length: 3579
Score: 1459.50 Matches: 370
Percent Similarity: 31.04% Conservative: 0
Best Local Similarity: 31.04% Mismatches: 3
Query Match: 76.41% Indels: 819
DB: 18 Gaps: 1
US-09-544-776-2 (1-373) x US-10-267-502-212 (1-3579)
QY 1 MetGluAspLeuAspGlnSerProIleuValSerSerSerAspSerProArgProGln 20
Db 1 ATGGAAGACCTGGACAGCTCTCTGTGCTCGTCTCTGGACAGCCCAACCCCGGCGCAG 60
QY 21 ProAlaPheLysTyrGlnPheValArgGluProGluAspGluGluGluGluGlu 40
Db 61 CCCGGCTTCAGTACAGTTCGTGAGGAGCCCGGAGGACGAGGAGAGAGAGAGGAG 120
QY 41 GluGluGluAspGluAspGluAspLeuGluGluLeuGluValLeuGluArgLysProAla 60
Db 121 GAAGAGGAGGACGAGGACGAGACCTGGAGGAGCTGGAGGTCTGTGAGAGGAAGCCCGC 180
QY 61 AlaGlyLeuSerAlaAlaProValProThrAlaProAlaAlaAlaGlyAlaProLeuMetAsp 80
Db 181 GCCGGGCTGTCCGCGGCCCCAGTGGCCCCCTGCGCGCGCGCGCGCTGATGGAC 240
QY 81 PheGlyAsnAspPheValProProAlaProArgGlyPheLeuProAlaAlaProVal 100
Db 241 TTCGGAATGACTTCTGTGCGCGCGCGCGCGCGCGCGCGCGCTCCCCCGCTC 300
QY 101 AlaProGluArgGlnProSerTrpAspProSerProValSerSerThrValProAlaPro 120
Db 301 GCCCGGAGCGGACGCGCTCTTGGGACCGAGCCCGGCTGTCTGTCGACCGTGC CGC 360
QY 121 SerPheLeuSerAlaAlaValSerProSerLysLeuProGluAspGluProPro 140
Db 361 TCCCCGCTGTCTGTGCGCGCAGTCTGCGCCTTCAAGCTCCCTGAGGACGACGAGCTCCG 420
QY 141 AlaArgProProProProProProAlaSerValSerProGlnAlaGluProValTrpThr 160
Db 421 GCCCGGCTTCCCCCTCTCCCCCGGCGCAGCGTGAGCCCCCAGGACAGAGCCCGTGTG 480
QY 161 ProProAlaProAlaProProSerThrProAlaAlaProLysArgArgGly 180
Db 481 CCGCCAGCCCGGCTTCCGCGCGCGCGCGCTCCACCCCGCGCGCGCCCAAGCGCAGGGGC 540
QY 181 SerSerGlySerVal----- 185
Db 541 TCCTCGGCTCAGTGGATGAGACCCCTTTTGTCTTCTGTCGATCTGAGCCTGTGATA 600
QY 185 ----- 185
Db 601 CGCTCCTCTGCAGAAAAATATGGACTTGAAGGAGCAGCAGGTAAACACTATTTCGGTGTG 660
QY 185 ----- 185
Db 661 CAAGAGGATTTCCCATCTGCTCTGTTGAACTGCTGCTTCTTCTTCTTCTTCTCTGTCCT 720
QY 185 ----- 185
Db 721 CTCTCAGCGCTTCTTTTCAAGAAACATGAATACCTTGGTGAATTTGTCAACAGTATTACCC 780
QY 185 ----- 185
Db 781 ACTGAAGGAACACTTCAAGAAAAATGTCAGTGAAGCTTCTTAAAGAGGTCTCAGAGAAGGCA 840
QY 185 ----- 185
Db 841 AAAAAGCTTACTCATAGATAGAGATTTAAACAGAGTTTTTTCAGAAATTAGAAATCTCAGAAATG 900
```

QY 185 ----- 185  
Db 901 GGATCATCGTTTCAGTGTCTCTCCAAAGCAGAACTCTGCCGTAAATAGTAGCAATCTCTAGG 960  
QY 185 ----- 185  
Db 961 GAAGAAATAATCGTGAATAATAAAGATGAAGAAGAGAGTGTAGTTAGTAATAACATCCTT 1020  
QY 185 ----- 185  
Db 1021 CATAATCAACAGAGTTACCTACAGCTCTTACTAAATTTGGTTAAAGAGATGAAGTTGTG 1080  
QY 185 ----- 185  
Db 1081 TCCTCAGAAAAGCAAAAGACAGTTTTTAATGAAAAGAGAGTTGCAGTGGAAAGCTCCTATG 1140  
QY 185 ----- 185  
Db 1141 AGGAGGAATATGCAGACTTCAAACCAATTTGAGCGAGTATGGGAAGTGAAGATAGTAAG 1200  
QY 185 ----- 185  
Db 1201 GAAGATAGTATATTTGGCTGTGGAGGTAAATCGAGAGCAACTTGGAAAGTAAAGTG 1260  
QY 185 ----- 185  
Db 1261 GATAAAAAATGTTTGCAGATAGCTTTGAGCAAACTAATCACGAAAAAGATAGTGAGAGT 1320  
QY 185 ----- 185  
Db 1321 AGTAATGATGATACTTCTTTCCAGTAGTACGCCAGAGGTATAAAGGATCGTCCAGGAGCA 1380  
QY 185 ----- 185  
Db 1381 TATATCACATGTGCTCCCTTTAACCCAGCAACTGAGAGCAATTCGAACAAACATTTTT 1440  
QY 185 ----- 185  
Db 1441 CCTTTGTTAGGAGATCCTACTTCAGAAAAATAAGACCGTGAAAAAATAAGAAAG 1500  
QY 185 ----- 185  
Db 1501 AAGGCCAAATAGTAACAGAGAGAATACTAGCACCAAAACATCAAAACCTTTTCTTGTA 1560  
QY 185 ----- 185  
Db 1561 GCAGCAGAGATTCTGAGACAGATTATGTCAACACAGATHAATTTAACAAAGGTGACTGAG 1620  
QY 185 ----- 185  
Db 1621 GAAGTCGTGCAAAACATGCTGAAGGCTGACTCCAGATTTAGTACAGGAAGCATGTGAA 1680  
QY 185 ----- 185  
Db 1681 AGTGAATTGAATGAAGTTACTGGTACAAAGATTGCTTATGAAACAAAAATGGACTTGGTT 1740  
QY 185 ----- 185  
Db 1741 CAAACATCAGAAGTTATGCAAGAGTCACTCTATCTCTGCAGCAGACTTTGGCCCATCATTT 1800  
QY 185 ----- 185  
Db 1801 GAAGAGTCAGAAGCTACTCCTTCACCAGTTTTCCTGACATTGTTATGGAAGCACCATTG 1860  
QY 185 ----- 185  
Db 1861 AATTCTGAGTTCTCTAGTGTGGTCTTCGGTGATACAGCCAGCTCATCACCATTAGAA 1920  
QY 185 ----- 185  
Db 1921 GCTTCTTCAGTTAATTATGAAGCATAAACATGAGCCTGAAAAACCCCCACCATTAGAA 1980

QY 185 ----- 185  
Db 1981 GAGGCCATGAGTGTATCACTTAAAAAAGTATCAGGAATAAAGGAAGAAATTAAGAGCCT 2040  
QY 185 ----- 185  
Db 2041 GAAATATTAAATGAGCTCTTCAAGAAACAGAGACTCCTTATATATCTATTGTCATGTGAT 2100  
QY 185 ----- 185  
Db 2101 TTAATTAAGAAACAAAGCTTTCTGCTGAACAGCTCCGGATTTCTCTGATTAATTCAGAA 2160  
QY 185 ----- 185  
Db 2161 ATGGCAAAAGTTGAACAGCCAGTGCCTGATCACTCTGAGCTAGTTGAAGATTCCCTCACCT 2220  
QY 185 ----- 185  
Db 2221 GATTCTGAACAGTTGACTTATTTTAGTGATGATTTCAATACCTGACGTTCCACAAAACAA 2280  
QY 185 ----- 185  
Db 2281 GATGAACTGTGATGCTTGTGAAGAAAGTCTCACTGAGACTTCAITTTGAGTCAATGATA 2340  
QY 185 ----- 185  
Db 2341 GAATATGAAAAATAAGGAAAAAATCACTGCTTTGCCACCTGAGGGAGGAAGCCATATTTG 2400  
QY 185 ----- 185  
Db 2401 GAATCTTTTAAGCTAGTTTAGATAACAAAAAGATACCTGTTACCTGATGAAGTTTCA 2460  
QY 185 ----- 185  
Db 2461 ACATTCGCAAAAAGAGGAAAAATTCCTTTGCAGATGGAGAGCTCAGTACTGCAGTTTAT 2520  
QY 185 ----- 185  
Db 2521 TCAATGATGACTTATTATTATTCTAAGGAAGCACAGATAAGAGAACTGAAAGCTTTTCA 2580  
QY 185 ----- 185  
Db 2581 GATTCAATCTCCAATTGAATATAGATGAGTTCCCTACATTTGATGATGATCTTCAAACTGAT 2640  
QY 185 ----- 185  
Db 2641 TCATTTTCTAAATTAGCCAGGGAATATACTGACCTAGAAGTATCCCAACAAAGTGAATTT 2700  
QY 185 ----- 185  
Db 2701 GCTAATGCCCGATGGAGCTGGGTCAATTTGCTTGCACAGAAATTGCCCATGACCTTTCT 2760  
QY 185 ----- 185  
Db 2761 TTGAAGAACATACAAACCCAAAGTTGAAGAGAAAAATCAGTTTCTCAGATGACTTTCTAAA 2820  
QY 185 ----- 185  
Db 2821 AATGGGTCTGCTACATCAAAAGGTGCTCTTATTGCTCTCAGATGTTTCTGCTTTGGCCACT 2880  
QY 185 ----- 185  
Db 2881 CAAGCAGAGATAGAGACATAGTTAAACCCAAAGTTCTTGTGAAGAAAGCTGAGAAAAAA 2940  
QY 185 ----- 185  
Db 2941 CTTCTCTCCGATACAGAAAAAGAGGACAGATCACCATCTGCTATATTTTCAGCAGAGCTG 3000  
QY 186 ----- 186  
Db 3001 AGTAAACTTCAGTTGTGCTGCTCTCTGTTACTGGAGAGACATTAAGAAAGCTGGAGTGGTG 3060  
QY 202 PheGlyAlaSerIleuPheLeuSerLeuThrValPheSerIleValSerValThr 221

3061	Db		TTTTGGTCCAGCCGCTATTTCCTGCTGCTTTCAATGTGACAGTATTTCAGCATTTGTGAGCGTAACA	3120
222	Qy		AlaTyrIleAlaLeuAlaLeuLeuSerValThrIleSerProArgIleTyrLysGlyVal	241
3121	Db		GCCTACATTTGCTTGGCCCTGCTCTCTGTGACCAATCAGCTTTAGGATATATACAGGGHGTG	3180
242	Qy		IleGlnAlaIleGlnLysSerAspGluGlyHisProPheArgAlaTyrLeuGluSerGlu	261
3181	Db		ATCCAAGCTATTCAGAAATCAGATGAAGGCCACCATTTCAGGCCATATCTGGAATCTGAA	3240
262	Qy		ValAlaIleSerGluGluLeuValGlnLysTyrSerAsnSerAlaLeuGlyHisValAsn	281
3241	Db		GTTTGCTATATCTGAGGAGTTGGTTCAAGAAGTACAGTAATTTCTGCTCTTGCTATGTGAAC	3300
282	Qy		CysThrIleLysGluLeuArgArgLeuPheLeuValAspAspLeuValAspSerLeuLys	301
3301	Db		TGCACGATTAAGGAACTCAGCGGCTCTTCTTAGTTGATGATTTAGTGTGATTTCTGTGAAG	3360
302	Qy		PheAlaValLeuMetTrpValPheThrTyrValGlyAlaLeuPheAsnGlyLeuThrLeu	321
3361	Db		TTTTCAGTGTGATGTGGTATTTTACCTATGTTGGTGCCTTGTTTAATGCTCTGACACTA	3420
322	Qy		IleIleLeuAlaIleIleSerLeuPheSerValProValIleTyrGluArgHisGlnAla	341
3421	Db		CTGATTTTGGCTCTCATTTCACTCTTCAGTGTGCTCTGTATTATTAATGAACGGCATCAGCGC	3480
342	Qy		GlnIleAspHisTyrLeuGlyLeuAlaAsnLysAsnValLysAspAlaMetAlaLysIle	361
3481	Db		CAGNTAGTCAATATCTAGGACTTGCAATTAAGNATGTTTAAGNATGCTATGGCTAAATC	3540
362	Qy		GlnAlaLysIleProGlyLeuLysArgLysAlaGlu	373
3541	Db		CAAGCAAAATTCCTCGATTGAACGCGAAAGCTGAA	3576

RESULT 13

```

US-10-327-213-8
; Sequence 8, Application US/10327213
; Publication No. US20040121341A1
; GENERAL INFORMATION:
; APPLICANT: FILBIN, MARIE T.
; APPLICANT: DOMENICONI, MARCO
; APPLICANT: CAO, ZIXUAN
; TITLE OF INVENTION: INHIBITORS OF MYELIN-ASSOCIATED GLYCOPROTEIN (MAG)
; FILE REFERENCE: CUNY/003
; CURRENT APPLICATION NUMBER: US/10/327,213
; CURRENT FILING DATE: 2002-12-20
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 3579
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-327-213-8

```

[illegible]

US-09-544-776-2 (1-373) x US-10-327-213-8 (1-3579)

Qy	1	MetGluAspLeuAspGlnSerProLeuValSerSerSerAspSerProProArgProGln	20
Db	1	ATGGAAGACCTGGACACAGTCTCTCTGTCTCGTCTCGACAGCCACCCCGCCGCGAG	60
Qy	21	ProAlaPheLysTyrGlnPheValArgGluProGluAspGluGluGluGluGluGlu	40
Db	61	CCCGCGTTCACGTACCAAGTTCTGAGGAGAGCCCGAGGACGAGGAGGAAGAGGAGGAG	120

185 ----- 185  
1201 GAAGATAGTGATATGTTGGCTGCTGGAGGTAAATCGAGAGCAACTGGAAAGTAAAGTG 1260  
185 ----- 185  
1261 GATAAAAATGTTTTCAGATAGCCTTGAGCAAACTAATCAGAAAAGATAGTGAGAT 1320  
185 ----- 185  
1321 AGTAATGATGATACTTCTTTCCCGTACCCAGTACCCAGAGGTATTAAGGATCGTCCAGGACA 1380  
185 ----- 185  
1381 TATATCACATGTGCTCCCTTTAAACCCAGCAGCAACTGAGAGCATTTGCAACAAACATTTTT 1440  
185 ----- 185  
1441 CCTTTGTTAGGAGATCCTACTTCAGAAAATAAGACCGATGAAAAAATAAGAGAAAG 1500  
185 ----- 185  
1501 AAGGCCAATAGTAACAGAGAGAGATACTAGCACCAAAACATCAAAACCTTTTCTTGTA 1560  
185 ----- 185  
1561 GCAGCACAGGATTTCTGAGACAGATTATGTCAACACAGATAATTTAAACAAAGGTGACTGAG 1620  
185 ----- 185  
1621 GAAGTCGTGCAAAACATGCTGGAAGCCTGACTCCAGATTTAGTA CAGGAAGCATGTGAA 1680  
185 ----- 185  
1681 AGTGAATTGAATGAGTTACTGTGTA CAAAGATTGCTTATGAACAAAATGGACTTGGTT 1740  
185 ----- 185  
1741 CAAACATCAGAAGTTATGCAAGAGTCACTCTATCCTGACGACAGCTTTGGCCCATCATTT 1800  
185 ----- 185  
1801 GAAGAGTCAGAAGCTACTCCTTACCAGTTTTCCTGACATTTGCTGACATTTGTTATGGAAGCACCATTG 1860  
185 ----- 185  
1861 AATTCTGCAGTTCTCTAGTGTGGTGCTTCGTGATACAGCCAGCTCATCACCATTAGAA 1920  
185 ----- 185  
1921 GCTTCTTCAGTTAATTATGAAGCATAAACATGAGCCTGAAAAACCCCCACCATATGAA 1980  
185 ----- 185  
1981 GAGGCCATGAGTGATACACTAAAAAAGTATCAGGAATAAGAGAAATTAAGAGCCT 2040  
185 ----- 185  
2041 GAAAATATTATGCAGCTCTTCAAGAAACAGAGCTCCTTATATATCTATTGTCATGTGAT 2100  
185 ----- 185  
2101 TTAATTAAAGAAACAAAGCTTTCTGCTGAACAGCTCCGGATTCTCTGTATTATTCAGAA 2160  
185 ----- 185  
2161 ATGGCAAAAGTTGAACAGCAGTGCCTGATCAATCTGAGCTAGTTGAAGATTCCTCACCT 2220  
185 ----- 185  
2221 GATTCTGAACCAAGTTGACTTATTATTAGTGATGATTCAATCCTGACGTTCCACAAAAACAA 2280  
185 ----- 185

2281 GATGAAACTGTGATGCTTGTGAAGAAAGTCTCACTGAGACTTCAATTGAGTCAATGATA 2340  
185 ----- 185  
2341 GAATATGAATAAAGGAAAAAATCTCAGTGCTTTGCCACCTGAGGGAGGAAAGCCATATTG 2400  
185 ----- 185  
2401 GAATCTTTTAAGCTCAGTTTAGATACACAAAAAGATACCCTGTTACCTGATGAAGTTTCA 2460  
185 ----- 185  
2461 ACATTGAGCAAAAAGGAGAAAAATTCCTTTGAGATGAGAGCTCAGTACTGAGTTTAT 2520  
185 ----- 185  
2521 TCAATGATGACTTATTATTATTCTAAGGAAGCACAGATAAGAGAAACTGAAACGTTTTCA 2580  
185 ----- 185  
2581 GATTCACTCCAATTGAAATTATAGATGAGTTCCCTACATTGATCAGTTCTTAAACCTGAT 2640  
185 ----- 185  
2641 TCATTTTCTAAATTAGCCAGGGAATATCTGACCTAGAAAGTATCCACAAAAGTGAAAT 2700  
185 ----- 185  
2701 GCTAATGCCCGATGGAGCTGGTCAATTGCTTGACAGAAATGGCCCATGACCTTTCT 2760  
185 ----- 185  
2761 TTGAAGAACATACAAACCCAAAGTTGAAGAGAAAAATCAGTTTCTCAGATGACTTTCTTAA 2820  
185 ----- 185  
2821 AATGGGTCTGTACATCAAAGGTGCTTTATTGCTCCAGATGTTTCTGCTTTGGCCACT 2880  
185 ----- 185  
2881 CAAGCAGATAGAGAGCATAGTTAAACCCAAAGTTCTTGTGAAAGAGCTGAGAAAAA 2940  
185 ----- 185  
2941 CTTCTCTCGATACAGAAAAAGAGACAGATCACCATCTGCTATATTTTTCAGCAGAGCTG 3000  
186 -----  
3001 AGTAAAACTTTCAGTTGTTGACCTCCTGTACTGGAGAGACATTAAAGAAAGCTGGAGTGGT 3060  
202 PheGlyAlaSerLeuPheLeuLeuSerLeuThrValPheSerIleValSerValThr 221  
3061 TTTGGTCAGAGCTATTCTCTGCTTTCATTGACAGATTAATCAGCAATTTGTGAGCGTAACA 3120  
222 AlaTyrIleAlaLeuAlaLeuSerValThrIleSerProargIleTyrIysGlyVal 241  
3121 GCCTACATTTGCTTGGCCCTGCTCTCTGTGACCATCAGCTTTAGGATATACAAGGGTGTG 3180  
242 IleGlnAlaIleGlnIysSerAspGluGlyHisProPheArgAlaTyrIleuGluSerGlu 261  
3181 ATCCAGCTATCCAGAAATCAGATGAAGGCCACCCCATCAGGGCATATCTGGAATCTGAA 3240  
262 ValAlaIleSerGluGluLeuValGlnIysTyrSerAsnSerAlaLeuGlyHisValAsn 281  
3241 GTTGTCTATCTGAGAGATTGGTTCAAGAGTACAGTAATCTGCTCTTGGTCATGTGAAC 3300  
282 CysThrIleIysGluLeuArgLeuPheLeuValAspAspLeuValAspSerLeuIys 301  
3301 TGCACGATAAAGGAACCTCAGGCCCTCTTCTTAGTTGATGATTTAGTTAGTTCTCTGAAG 3360  
302 PheAlaValLeuMetTrpValPheThrTyrValGlyAlaLeuPheAsnGlyLeuThrLeu 321

Db 3361 TTTCAGTGTGTGATGGGTATTTACCTATGTTGGTGCCTGTTTAAATGGTCTGCACACTA 1420  
Qy 322 LeuileuAlaLeuileuSerLeuPheSerValProValIleTyRGluArgHisGlnAla 341  
Db 3421 CTGATTTTGGCTCTCATTTTCACTCTTCAGTGTTCCTGTTATTATGAACGGCATCAGGGG 3480  
Qy 342 GlnileAspHisTyRLeuGlyLeuAlaAsnLysAsnValLysAspAlaMetAlaLysIle 361  
Db 3481 CAGATAGATCATTTATCTAGACTTTGCAATAGAATGTTAAAGATGCTATGGCTAAATC 3540  
Qy 362 GlnAlaLysIleProGlyLeuLysArgLysAlaGlu 373  
Db 3541 CAAGCAAAAATCCCTGGATTGAAGCGCAAAAGCTGAA 3576

## RESULT 14

US-10-466-258-8  
; Sequence 8, Application US/10466258  
; Publication No. US20040132096A1  
; GENERAL INFORMATION:  
; APPLICANT: GLAXO GROUP LIMITED  
; TITLE OF INVENTION: ASSAY  
; FILE REFERENCE: P80966 GCW  
; CURRENT APPLICATION NUMBER: US/10/466,258  
; CURRENT FILING DATE: 2003-07-15  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 8  
; LENGTH: 3579  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(3579)  
US-10-466-258-8

Alignment Scores:  
Pred. No.: 1,38e-120 Length: 3579  
Score: 1459.50 Matches: 370  
Percent Similarity: 31.04% Conservative: 0  
Best Local Similarity: 31.04% Mismatches: 3  
Query Match: 76.41% Indels: 819  
DB: 19 Gaps: 1

US-09-544-776-2 (1-373) x US-10-466-258-8 (1-3579)

Qy 1 MetGluAspLeuAspGlnSerProLeuValSerSerAspSerProProArgProGln 20  
Db 1 ATGGAAGACCTGGACCAGTCTCCTCTGGTCTCGTCTCGGACAGGCCACCCCGGCGCAG 60  
Qy 21 ProAlaPheLysTyRGlnPheValArgGluProGluAspGluGluGluGlu 40  
Db 61 CCCGCGTTCAAGTACCAGTTCGTGGAGGAGCCGAGGACGAGGAGGAAGAGGAGGAG 120  
Qy 41 GluGluGluAspGluAspGluLeuGluGluLeuGluValLeuGluArgLysProAla 60  
Db 121 GAAGAGGAGGACGAGGACGACGACCTTGAGGAGCTGGAGGTGCTGGAGAGGAGCCCGCC 180  
Qy 61 AlaGlyLeuSerAlaAlaProValProThrAlaProAlaProAlaGlyAlaProLeuMetAsp 80  
Db 181 GCCGGGCTGTCCGCGGCCCGAGTGGCCACCGCCCTTGCCTGCGCGGCGCGCCCTGATGGAC 240  
Qy 81 PheGlyAsnAspPheValProProAlaProArgGlyPheLeuProAlaAlaProProVal 100  
Db 241 TTCGGAATGACTTGTGCGCGCGCGCGCGCGCGGACCCCTGCGCGCCCTCCCCCGGTC 300  
Qy 101 AlaProGluArgGlnProSerTrpAspProSerProValSerSerThrValProAlaPro 120  
Db 301 GCCCGGAGCGGACGCGCTTGGGACCCGAGCCCGGTGCTGCTGACCGGTGCCCGGCCA 360  
Qy 121 SerPheLeuSerAlaAlaValSerProSerLysLeuProGluAspAspGluProPro 140  
Db 361 TCCCGCTGTCTGTCGCGCAGTCTCGCCCTCCAAGCTCCCTGAGGACGACGAGCCTCGG 420

Qy 141 AlaArgProProProProProAlaSerValSerProGlnAlaGluProValTrpThr 160  
Db 421 GCCCGGCTCCCCCTCTCCCGGCGGAGGTGAGCCCGGAGAGAGCCCGGTGGACC 480  
Qy 161 ProProAlaProAlaProAlaProProSerThrProAlaAlaProLysArgArgGly 180  
Db 481 CGCCAGCCCGGCTCCCGCGCGCCCTCCACCCCGCGCGCCCAAGCGCAGGGGC 540  
Qy 181 SerSerGlySerVal----- 185  
Db 541 TCCTCGGCTCAGTGGATGAGACCCCTTTTGTCTTCTCTGCTGTCATCTGAGCCTGTGATA 600  
Qy 185 ----- 185  
Db 601 CGCTCCTCTGCAGAAAAATATGGACTTTGAAGGAGCAGCCAGGTAACTATTTTCGGCTGT 660  
Qy 185 ----- 185  
Db 661 CAAGAGGATTTCCCATCTGTCTGTCTGAACTGCTGCTTCTCTTCTCTCTCTCTCTCT 720  
Qy 185 ----- 185  
Db 721 CTCTCAGCGCTTCTTTCAAAGAACATGAATACCTTGGTAAATTTGTCAACAGTATTACCC 780  
Qy 185 ----- 185  
Db 781 ACTGAAGGAACACTTCAAGAAAAATGTCAGTGAGCTTCTTAAAGAGGTCTCAGAGAGGCA 840  
Qy 185 ----- 185  
Db 841 AAAACTCTACTCATAGATAGAGATTTAACAGAGTTTTTCAAGATTTAGAAATCTCAGAAATG 900  
Qy 185 ----- 185  
Db 901 GGATCATCGTTTCAGTGTCTCTCAAAGCAGAACTGCCGTAATAGTAGCAAACTCTAGG 960  
Qy 185 ----- 185  
Db 961 GAAGAAATAATCGTGAAAAATAAAGATGAAGNAGAGAGTAGTTAGTTAATAACATCCTT 1020  
Qy 185 ----- 185  
Db 1021 CATAATCAACAAGAGTTTACTACAGCTCTTACTAAATTTGGTTAAAGAGGATGAAGTTGTG 1080  
Qy 185 ----- 185  
Db 1081 TCTTCAGAAAAAGCAAAAGACAGTTTTTAATGAAAAGAGAGTTGTCAGTGGAGCTCCTATG 1140  
Qy 185 ----- 185  
Db 1141 AGGGAGGAATATGCAGACTTCAAACCATTTGAGCGAGTATGGGAAGTGAAGATAGTAAG 1200  
Qy 185 ----- 185  
Db 1201 GAAGATAGTATATGTGTGGTCTGCGAGTAAATCGAGAGCAACTTGGAAAGTAAAGTG 1260  
Qy 185 ----- 185  
Db 1261 GATAAAAAATGTTTTCAGATAGACCTTGAGCAAACTAATACAGAAAAAGATAGTAGAGT 1320  
Qy 185 ----- 185  
Db 1321 AGTAATGATGATCTTCTTTCCCAAGTACGCCAGAGGTATATAAGATCGTCCAGGACA 1380  
Qy 185 ----- 185  
Db 1381 TATATCATGTGTCTCCCTTTAACCCAGCAGCAACTGAGAGCATTTGCAACAACTTTTT 1440  
Qy 185 ----- 185  
Db 1441 CCTTTGTTAGGAGATCTTACTTTTCAAAAAATAGACCGATGAAAAAATAATAGAGAAAG 1500  
Qy 185 ----- 185



Db	1501	AAGGCCCAAAATAGTAA	CAGAGAAGAAATAC	TAGCACCAAAACATCAAAC	CTTTCTTGTA	1560
Qy	185	-----	-----	-----	-----	185
Db	1561	GCAGCACAGGATTCTG	GAGACAGATTATGT	CACAACAGATAATTTAA	CAAAAGGTGACTGAG	1620
Qy	185	-----	-----	-----	-----	185
Db	1621	GAAGTCGTGGCAACAT	AGCCTGAGAGCCTGACT	CCAGATTTAGTACAGAGCATGTGA	1680	
Qy	185	-----	-----	-----	-----	185
Db	1681	AGTGAATTGAATGAAG	TACTGGTACAAAGATTGCT	TATGAAACAAAATGGACTTGGTT	1740	
Qy	185	-----	-----	-----	-----	185
Db	1741	CAAAATCATGAAAGTT	ATGCAAGAGTCACTCT	ATCTCTGCACACAGCTTTG	CCCATCATTT	1800
Qy	185	-----	-----	-----	-----	185
Db	1801	GAAGAGTCAGAAGCTA	CTCCTTACCAGTTTTTG	CTGCATTTGTTATGAGACCACTG	1860	
Qy	185	-----	-----	-----	-----	185
Db	1861	AATTCTGCAGTTCTAG	TGCTGGTTCGGTGATAC	AGCCAGCTCATCACCATTAGAA	1920	
Qy	185	-----	-----	-----	-----	185
Db	1921	GCTTCTTCAGTTAATT	ATGAAAGCATAAACAT	GAGCCTGAAAACCCCCACCATATGA	1980	
Qy	185	-----	-----	-----	-----	185
Db	1981	GAGGCCATGAGTGTAT	CACTAAAAAAGTATCAGGA	TAAAGGAAGAAATTAAAGAGCCT	2040	
Qy	185	-----	-----	-----	-----	185
Db	2041	GAANAATTAAATGCA	GTCTTCAAGAAA	CAGAAGCTCTTATATATCTATTG	CATGTGAT	2100
Qy	185	-----	-----	-----	-----	185
Db	2101	TTAATTAAGAAACAA	AGCTTTCTGTGAAC	CAGCTCCGGATTTCTGTGATTATCAGAA	2160	
Qy	185	-----	-----	-----	-----	185
Db	2161	ATGGCAAAAGTTGAA	CAGCCAGTCCGTGAT	CTCTGAGTAGTTGAAGATTCCTCACCT	2220	
Qy	185	-----	-----	-----	-----	185
Db	2221	GATTCTGAACCAAGT	TGACTTATTTAGTGATGATTCA	TACTACCTGAGCTTCCACAAAACAA	2280	
Qy	185	-----	-----	-----	-----	185
Db	2281	GATGAAACTGTGATG	CTTGTGAAAGAAAGTCT	CACCTGAGACTTCAITTGAGTCAATGATA	2340	
Qy	185	-----	-----	-----	-----	185
Db	2341	GAATATGAAAAATAG	GAAAAAAGTCTGTTGCC	ACTGAGGAGAAAGCCATATTG	2400	
Qy	185	-----	-----	-----	-----	185
Db	2401	GAATCTTTTAAGCTC	AGTTTAGATACACAAA	AGATACCTGTTCCTGATGAGATTTC	2460	
Qy	185	-----	-----	-----	-----	185
Db	2461	ACATTGAGCAAAAGG	AGAAAAATTCCTTTG	CAGATGGAGGAGCTCAGTACTGCAGTTTAT	2520	
Qy	185	-----	-----	-----	-----	185
Db	2521	TCAAATGATGACTTAT	TTTATTTCTAAGGA	AGCACAGATAGAGAAAC	TGAAACGTTTTCA	2580
Qy	185	-----	-----	-----	-----	185

Db	2581	GATTCACTCCAAATTGAAATTATAGATGAGTTCCTCACTATGATCACTTCTAAAACTGAT	2640
Qy	185	-----	185
Db	2641	TCATTTTCTAAATTAGCCAGGAATATACTGACCTAGAACTATCCACAAAAGTGAATTT	2700
Qy	185	-----	185
Db	2701	GCTAAATCCCGGATGGAGCTGGGTCAATTCGCTTGCCAGAGAAATGCCCCATGACCTTTCT	2760
Qy	185	-----	185
Db	2761	TTGAAGAACATACAAACCAAAAGTTGAAGAGAAAATCAGTTTCTCAGATGACTTTTCTAAA	2820
Qy	185	-----	185
Db	2821	AATGGGTCTGTACATCAAAGGTGCTCTTATTGCTCTCCAGATGTTTCTGCTTTGGCCACT	2880
Qy	185	-----	185
Db	2881	CAAGCAGAGATAGAGAGCATAGTTAAACCCAAAGTTCTTGTGAAGAAGCTGAGAAAAA	2940
Qy	185	-----	185
Db	2941	CTTCTTCCGATACAGAAAAAGGACAGATCACCATCTCTATATTTTCAGCAGAGCTG	3000
Qy	186	-----ValValAspLeuLeuTyrTrpArgAspIleIysIysThrGlyValVal	201
Db	3001	AGTAAAACTTTCAGTTGTTGACCTCTGTACTGGAGAGACATTAAAGAAGCTGGAGTGGTG	3060
Qy	202	PheGlyAlaSerLeuPheLeuLeuLeuSerLeuThrValPheSerIleValSerValThr	221
Db	3061	TTTGGTGCCAGCCTATTCTCTGCTGCTTTCATTGACAGTATTCAGCATTTGAGCGCTAAACA	3120
Qy	222	AlaTyrIleAlaLeuAlaLeuSerValThrIleSerProArgIleTyrIysGlyVal	241
Db	3121	GCCTACATTGCCTTGGCCCTGCTCTCTGTGACCATTGAGCTTTAGGATATACAGGGGTGTG	3180
Qy	242	IleGlnAlaIleGlnIysSerAspGluGlyHisProPheArgAlaTyrLeuGluSerGlu	261
Db	3181	ATCCAAGCTATCCAGAAATCAGATGAAGGCCACCCATTTCAGGGCATATCTGGAAATCTGAA	3240
Qy	262	ValAlaIleSerGluGluLeuValGlnIysTyrSerAsnSerAlaLeuGlyHisValAsn	281
Db	3241	GTTGCTATATCTCAGGAGTTGGTTCAGAAGTACAGTAATTCCTGCTCTTGCTCATGTGAAC	3300
Qy	282	CysThrIleIysGluLeuArgArgLeuPheLeuValAspAspLeuValAspSerLeuIys	301
Db	3301	TGCACGATAAAGNACTCAGGCCCTCTCTTCTAGTTGATGATTTAGTTGATTTCTCTGAAG	3360
Qy	302	PheAlaValLeuMetTrpValPheThrTyrValGlyAlaLeuPheAsnGlyLeuThrLeu	321
Db	3361	TTTGCAGTGTGATGTGGGTATTACCTATGTTGGTGCCTTGTTTAATGGTCTCGACACTA	3420
Qy	322	LeuIleLeuAlaLeuIleSerLeuPheSerValproValIleTyrGluArgHisGlnAla	341
Db	3421	CTGATTTTGCTCTCATTTTCACCTCTTCAGTGTTTCCCTGTTATTATTGAACGGCATCAGCG	3480
Qy	342	GlnIleAspHisTyrIysLeuGlyLeuAlaAsnIysAsnValIysAspAlaMetAlaIysIle	361
Db	3481	CAGATAGATCATTTCTAGGACTTGCAAATAAGAATGTTAAAGATGCTATGGCTAAAATC	3540
Qy	362	GlnAlaIysIleProGlyLeuIysArgIysAlaGlu	373
Db	3541	CAAGCAAAAATCCCTGGATTGAAGCGCAAAAGCTGAA	3576

RESULT 15  
US-10-810-653-22  
; Sequence 22, Application US/10810653  
; Publication No. US20040253218A1  
; GENERAL INFORMATION:  
; APPLICANT: EISENBACH-SCHWARTZ, Michael  
; APPLICANT: COHEN, Irun R.

; APPLICANT: BESERMAN, Pierre  
; APPLICANT: MOSONEGO, Alon  
; APPLICANT: MOALEM, Gila  
; TITLE OF INVENTION: ACTIVATED T-CELLS, NERVOUS SYSTEM-SPECIFIC ANTIGENS AND THEIR USE  
; FILE REFERENCE: EIS-SCHWARTZ=2A  
; CURRENT APPLICATION NUMBER: US/10/810,653  
; CURRENT FILING DATE: 2004-03-29  
; PRIOR APPLICATION NUMBER: US/09/893,348  
; PRIOR FILING DATE: 2001-06-28  
; PRIOR APPLICATION NUMBER: US 09/314,161  
; PRIOR FILING DATE: 1999-05-19  
; PRIOR APPLICATION NUMBER: US 09/218,277  
; PRIOR FILING DATE: 1998-12-22  
; PRIOR APPLICATION NUMBER: PCT/US98/14715  
; PRIOR FILING DATE: 1998-07-21  
; PRIOR APPLICATION NUMBER: IL 124500  
; PRIOR FILING DATE: 1998-05-19  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 22  
; LENGTH: 3579  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(3579)  
; OTHER INFORMATION:  
US-10-810-653-22

## Alignment Scores:

Pred. No.: 1 38e-120 Length: 3579  
Score: 1459.50 Matches: 370  
Percent Similarity: 31.04% Conservative: 0  
Best Local Similarity: 31.04% Mismatches: 3  
Query Match: 76.41% Indels: 819  
DB: 20 Gaps: 1

US-09-544-776-2 (1-373) x US-10-810-653-22 (1-3579)

QY 1 MetGluAspLeuaspGlnSerProLeuValSerSerSerAspSerProProArgProGln 20  
DB 1 ATGGAAGACCTGGACCAAGTCTCTCTGGTCTCGTCTCGACAGCCACCCCGCGCGCAG 60  
QY 21 ProAlaPheIysTyrGlnPheValArgGluProGluAspGluGluGluGluGlu 40  
DB 61 CCCGCGTTCAAGTACCAGTTCTGTAGGAGCCCGAGGACGAGGAGGAGGAGGAG 120  
QY 41 GluGluAspGluAspGluAspLeuGluGluLeuGluValLeuGluArgLysProAla 60  
DB 121 GAAGAGGAGGACGAGGACCAAGACCTGGAGGAGCTGGAGGTGCTGGAGAGGAGCCCGCC 180  
QY 61 AlaGlyLeuSerAlaAlaProValProThrAlaProAlaProAlaAlaGlyAlaProLeuMetAsp 80  
DB 181 GCCCGGCTGTCCGCGGCCCGAGTGTGCCACCGCCCTGCCGCGCGCGCGCTGTGATGGAC 240  
QY 81 PheGlyAsnAspPheValProProAlaProArgGlyPheLeuProAlaAlaProProVal 100  
DB 241 TTCGGAATGACTTCTGTCGCGCGGCCCGGCCCGCGGACCCCTGCCGCGCGCTCCCGCCGTC 300  
QY 101 AlaProGluArgGlnProSerTrpAspProSerProValSerSerThrValProAlaPro 120  
DB 301 GCCCGGAGCGGACCGCTTTGGGACCCGAGCCCGGTGCTGCGACCGCTGCCCGGCCA 360  
QY 121 SerPheLeuSerAlaAlaValSerProSerLysLeuProGluAspAspGluProPro 140  
DB 361 TCCCGCTGTCTGCTCCCGCAGTCTCGCCCTCCAAGCTCCCTGAGGACGACGAGCCTCG 420  
QY 141 AlaArgProProProProProProAlaSerValSerProGlnAlaGluProValTrpThr 160  
DB 421 GCCCGGCTTCCCGCTCTCCCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480  
QY 161 ProProAlaProAlaProAlaProProSerThrProAlaProLysArgArgGly 180

DB 481 CGCCAGCCCCGGCTCCCGCGCGCCCTCCACCCCGCGCGCCCAAGCGCAGGGC 540  
QY 181 SerSerGlySerVal----- 185  
DB 541 TCCTCGGGCTCAGTGGATGAGACCCCTTTTGTCTCTTCTCTGCTGCATCTGAGCCTGTGATA 600  
QY 185 ----- 185  
DB 601 CGCTCCTCTGCAGAAAAATATGGACTTGAAGGAGCAGCCAGGTAAACACTATTTCGGCTGT 660  
QY 185 ----- 185  
DB 661 CAAGAGGATTTCCCATCTGCTGCTTGAACCTGCTGCTTCTCTCTCTCTCTCTCTCT 720  
QY 185 ----- 185  
DB 721 CTCTCAGCGCTTCTTTCAAAGAACAATGAATACCTTGGTAAATTTGTCAACAGTATTACCC 780  
QY 185 ----- 185  
DB 781 ACTGAAGGAACACTTCAAGAAAAATGTCACTGAAGCTTTAAAGAGGTCTCAGAGAAGGCA 840  
QY 185 ----- 185  
DB 841 AAAACTCTACTCATAGATAGAGATTTTAAACAGAGTTTTTCAAGATTTAGATACTCAGAAATG 900  
QY 185 ----- 185  
DB 901 GGATCATCGTTCAGTGTCTCTCCAAAAGCAGAAATCTGCCGTAATAGTAGCAAAATCCTAGG 960  
QY 185 ----- 185  
DB 961 GAAGAAATTAATCGTGAAAAATAAAGATGAAGAAGAGAGTTAGTTAGTTAATACATCCTT 1020  
QY 185 ----- 185  
DB 1021 CATAATCAACAAGAGTTACCTACAGCTCTTAATAATTGGTTAAAGAGGATGAAGTTGTG 1080  
QY 185 ----- 185  
DB 1081 TCTTCAGAAAAAGCAAAAGACAGTTTTAAATGAAAAGAGAGTTGCAGTGGAACTCCTATG 1140  
QY 185 ----- 185  
DB 1141 AGGGAGGATATGCAGACTTCAACCATTTGAGCGGATATGGGAAGTGAAGTAGTAGTAAG 1200  
QY 185 ----- 185  
DB 1201 GAAGATAGTGATATGTTGGCTGCTGGAGGTAAATCGAGAGCAACTTGGAAAGTAAAGTG 1260  
QY 185 ----- 185  
DB 1261 GATAAAAAATGTTTTCAGATAGCCTTGAAGCAAACTAATACGAAAAAGATAGTGAGAGT 1320  
QY 185 ----- 185  
DB 1321 AGTAATGATGATCTTCTTTCCCCAGTAGCCAGGAAGGTATAAAGATCGTCCAGGAGCA 1380  
QY 185 ----- 185  
DB 1381 TATATCATGTGTCTCCCTTTAACCCAGCAGCAACTGAGAGCATTGCAACAAACATTTTT 1440  
QY 185 ----- 185  
DB 1441 CCTTTGTTAGGAGATCCTACTTTCAGAAAAATAAGCCGATGAAAAAATAAGAGAAGAG 1500  
QY 185 ----- 185  
DB 1501 AAGGCCCAATAGTAACAGAGAGAGATACTAGCACCAAAACATCAAACTTTTCTTTGTA 1560  
QY 185 ----- 185  
DB 1561 GCAGCACAGGATTTCTGAGACAGATTATGTCAACACAGATTATGTCAACAAAGGTGAGT 1620

QY 185 ----- 185  
Db 1621 GAAGTCGTGCAACATGCCTGAAGGCTGACTCCAGATTAGTACAGGAAGCATGTGAA 1680  
QY 185 ----- 185  
Db 1681 AGTGAATTGAATGAAGTTACTGGTACAAAGATTGCTTTATGAACAAAAATGGACTTGGTT 1740  
QY 185 ----- 185  
Db 1741 CAAACATCAGAAGTTATGCAAGAGTCACTCTATCCTGCAGCAGAGCTTTGCCCATCATTT 1800  
QY 185 ----- 185  
Db 1801 GAAGAGTCAGAAGCTACTCCTTCACCAAGTTTTCGCTGACATTGTTATGGAAGCACCATTG 1860  
QY 185 ----- 185  
Db 1861 AATTCTGCAGTTCTTAGTCTGGTCTTCGCTGATACAGCCAGCTCATCACCATTAGAA 1920  
QY 185 ----- 185  
Db 1921 GCTTCTTCAGTTAATTATGAAGCATAAACATGAGCCTGAAAAACCCCCACCATATGAA 1980  
QY 185 ----- 185  
Db 1981 GAGGCCATGAGTGTATCACTAAAAAAGATATCAGGAATAAGGAGAAATTAAGAGCCT 2040  
QY 185 ----- 185  
Db 2041 GAAAAATATTAAATGCGCTCTTCAAGAAACAGAGCTCCTTATATATCTATTGTCATGTGAT 2100  
QY 185 ----- 185  
Db 2101 TTAATTAAAGAAACAAAGCTTCTGCTGAACAGCTCCGGATTCTCTGATTAATTCAGAA 2160  
QY 185 ----- 185  
Db 2161 ATGCCAAAGTTGAAACAGCAGCGCTGATCATTTCTGAGCTAGTTGAAGATTCCTCACCT 2220  
QY 185 ----- 185  
Db 2221 GATTCTGAACCAAGTTGACTTATTATTAGTGATGATTCAATACCTGACGTTCCACAAAAACAA 2280  
QY 185 ----- 185  
Db 2281 GATGAACCTGTGATGCTTGTGAAGAAAGTCTCACTGAGACTTCATTTGAGTCAATGATA 2340  
QY 185 ----- 185  
Db 2341 GAATATGAAAAATAAGGAAAAACTCAGTGTCTTCCACCTGAGGGAGGAAAGCCATATTG 2400  
QY 185 ----- 185  
Db 2401 GAATCTTTAAGCTCAGTTTAGATAACACAAAAAGATACCCCTGTTACCTGATGAAGTTTCA 2460  
QY 185 ----- 185  
Db 2461 ACATTGAGCAAAAAGGAGAAAAATTCCTTTGCAGATGAGGAGCTCAGTACTGCAGTTTAT 2520  
QY 185 ----- 185  
Db 2521 TCAATGATGACTTATTATTATTCTAAGGAAGCAGACATAGAGAACTGAAACGTTTCA 2580  
QY 185 ----- 185  
Db 2581 GATTTCATCCAAATGAAATTATAGATGAGTTCCCTACATTGATCAGTCTCAAAAACGTGAT 2640  
QY 185 ----- 185  
Db 2641 TCATTTTCTAAATTAGCCAGGGAATATACCTGACCTAGAGATATCCCAACAAAGTGAAATT 2700

QY 185 ----- 185  
Db 2701 GCTAATGCCCGGATGGAGCTGGGTCAATTGGCTTGCACAGAAATTGCCCATGACCTTTCT 2760  
QY 185 ----- 185  
Db 2761 TTGAAGAACATACAAACCCAAAGTTGAAGAGAAAAATCAGTTTCTCAGATGACTTTTCTAAA 2820  
QY 185 ----- 185  
Db 2821 AATGGGTCTGTACTACATAAAGGTGCTCTTATTGCCCTCCAGATGTTTCTGCTTTGGCCACT 2880  
QY 185 ----- 185  
Db 2881 CAAGCAGAGATAGAGAGCATAGTTAAACCCAAAGTTCTTGTGAAGAAGCTGAGAAAAA 2940  
QY 185 ----- 185  
Db 2941 CTTCTTCCGATACAGAAAAAGAGCAGATCACCATCTGCTATATTTTTCAGCAGAGCTG 3000  
QY 186 -----ValValAspLeuTyrTrpArgAspIleLysLysThrGlyValVal 201  
Db 3001 AGTAAAACTTCAGTTGTTGACCTCTGTACTGGAGAGACATTAAAGAAGACTCGAGTGGTG 3060  
QY 202 PheGlyAlaSerLeuPheLeuSerLeuThrValPheSerIleValSerValThr 221  
Db 3061 TTTGGTGCAGGCTATTCTCTGCTTTCATTGACAGTATTCAGCATTTGTGAGCGTAACA 3120  
QY 222 AlaTyrIleAlaLeuAlaLeuSerValThrIleSerProArgIleTyrLysGlyVal 241  
Db 3121 GCTCATTTGCTTGGCCCTGCTCTCTGTGACCATCAGCTTTTAGGATATACAAAGGTGTG 3180  
QY 242 IleGlnAlaIleGlnLysSerAspGluGlyHisProPheArgAlaTyrLeuGluSerGlu 261  
Db 3181 ATCCAGCTATCCAGAAATCAGATGAAGGCCACCCATTCAAGGCATATCTGGAATCTGAA 3240  
QY 262 ValAlaIleSerGluGluLeuValGlnLysTyrSerAsnSerAlaLeuGlyHisValAsn 281  
Db 3241 GTTGCTATATCTGAGGAGTTGGTTCAAGAAGTACAGTAATTCGTCTTGTGTCATGTGAAC 3300  
QY 282 CysThrIleLysGluLeuLeuArgLeuPheLeuValAspLeuValAspSerLeuLys 301  
Db 3301 TGACGATAAAGAACTCAGGCGCCTCTTCTTAGTTGATGATTAGTTGATTCTCTGAAG 3360  
QY 302 PheAlaValLeuMetTrpValPheThrTyrValGlyAlaLeuPheAsnGlyLeuThrLeu 321  
Db 3361 TTTGCAAGTGTGATGTTGGGTATTTACTATGTTGGTGCCTTGTATTAATGGTCTGACACTA 3420  
QY 322 LeuIleLeuAlaLeuLysSerLeuPheSerValProValIleTyrGluArgHisGlnAla 341  
Db 3421 CTGATTTTGGCTCTCAATTTCACTTTCAGTGTTCCTTGTATTATTAATGAACGGCATCAGCG 3480  
QY 342 GlnIleAspHisTyrLeuGlyLeuAlaAsnLysAsnValLysAspAlaMetAlaLysIle 361  
Db 3481 CAGATAGATCATTTATCTAGGACTTGCAATAAGAAATGTTAAAGATGCTATGGCTAAATC 3540  
QY 362 GlnAlaLysIleProGlyLeuLysArgLysAlaGlu 373  
Db 3541 CAAGCAAAAATCCCTGGATTGAAGCCAAAGCTGAA 3576

Search completed: August 3, 2005, 23:37:50  
Job time : 743 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 3, 2005, 17:11:04 ; Search time 1407 Seconds  
(without alignments)

10304.249 Million cell updates/sec

Title: US-09-544-776-1

Perfect score: 2240

Sequence: 1 cgtcacacagtaggtccctc.....taaaaaaaaaaaaaaaaaa 2240

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 7287783 seqs, 3236178273 residues

Total number of hits satisfying chosen parameters: 8279544

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:\*

1: /cgn2\_6/ptodata/2/pubpna/PCT NEW PUB.seq.\*  
2: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW PUB.seq.\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW PUB.seq.\*  
6: /cgn2\_6/ptodata/2/pubpna/PCRU5\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW PUB.seq.\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW PUB.seq.\*  
13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/2/pubpna/US10D\_PUBCOMB.seq.\*  
17: /cgn2\_6/ptodata/2/pubpna/US10E\_PUBCOMB.seq.\*  
18: /cgn2\_6/ptodata/2/pubpna/US10F\_PUBCOMB.seq.\*  
19: /cgn2\_6/ptodata/2/pubpna/US10G\_PUBCOMB.seq.\*  
20: /cgn2\_6/ptodata/2/pubpna/US10H\_PUBCOMB.seq.\*  
21: /cgn2\_6/ptodata/2/pubpna/US10I\_PUBCOMB.seq.\*  
22: /cgn2\_6/ptodata/2/pubpna/US10\_NEW PUB.seq.\*  
23: /cgn2\_6/ptodata/2/pubpna/US11A\_PUBCOMB.seq.\*  
24: /cgn2\_6/ptodata/2/pubpna/US11\_NEW PUB.seq.\*  
25: /cgn2\_6/ptodata/2/pubpna/US60\_NEW PUB.seq.\*  
26: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	50	2.2	50	17	US-10-131-827-743
2	37.4	1.7	48	16	US-10-275-071-34
3	37.4	1.7	50	16	US-10-275-071-17
4	35.2	1.6	48	9	US-09-962-055-36
5	35.2	1.6	48	9	US-09-976-740-36
6	35.2	1.6	48	13	US-10-023-529-36
7	35.2	1.6	48	13	US-10-023-523-36

8	35.2	1.6	48	17	US-10-616-187-36	Sequence 36, Appl
9	35.2	1.6	48	18	US-10-671-242-36	Sequence 36, Appl
10	29.6	1.3	46	21	US-10-820-487-9	Sequence 9, Appl
11	29.6	1.3	46	21	US-10-903-975-9	Sequence 9, Appl
12	28.6	1.3	36	17	US-10-418-182-86	Sequence 86, Appl
13	28.6	1.3	36	21	US-10-479-472A-12	Sequence 12, Appl
14	28	1.2	45	9	US-09-962-055-30	Sequence 30, Appl
15	28	1.2	45	9	US-09-976-740-30	Sequence 30, Appl
16	28	1.2	45	13	US-10-023-529-30	Sequence 30, Appl
17	28	1.2	45	13	US-10-023-523-30	Sequence 30, Appl
18	28	1.2	45	17	US-10-616-187-30	Sequence 30, Appl
19	28	1.2	45	18	US-10-671-242-30	Sequence 30, Appl
20	27	1.2	31	9	US-09-801-274-331	Sequence 331, App
C 21	26.4	1.2	36	17	US-10-418-182-84	Sequence 84, Appl
C 22	26.4	1.2	36	19	US-10-296-085A-74	Sequence 74, Appl
C 23	25.8	1.2	30	13	US-10-101-487-2	Sequence 2, Appl
C 24	25.8	1.2	30	13	US-10-101-487-65	Sequence 65, Appl
C 25	25.8	1.2	30	21	US-10-939-988-2	Sequence 2, Appl
C 26	25.8	1.2	30	21	US-10-939-988-65	Sequence 65, Appl
C 27	25.6	1.1	47	16	US-10-315-515-112	Sequence 112, App
28	25.6	1.1	50	8	US-08-781-986A-2153	Sequence 2153, Ap
29	25.6	1.1	50	18	US-10-329-624-2153	Sequence 2153, Ap
30	25.4	1.1	40	16	US-10-187-253A-54	Sequence 54, Appl
31	25.2	1.1	30	13	US-10-101-487-1	Sequence 1, Appl
32	25.2	1.1	30	13	US-10-101-487-63	Sequence 63, Appl
C 33	25.2	1.1	30	14	US-10-057-467-12	Sequence 12, Appl
C 34	25.2	1.1	30	21	US-10-939-988-1	Sequence 1, Appl
35	25.2	1.1	30	21	US-10-939-988-63	Sequence 63, Appl
36	25.2	1.1	50	20	US-09-306-333A-27	Sequence 27, Appl
37	25	1.1	25	19	US-10-717-597-4349	Sequence 4349, Ap
38	25	1.1	25	19	US-10-717-597-4350	Sequence 4350, Ap
39	25	1.1	25	19	US-10-717-597-4351	Sequence 4351, Ap
40	25	1.1	25	19	US-10-717-597-4352	Sequence 4352, Ap
41	25	1.1	25	19	US-10-717-597-4353	Sequence 4353, Ap
42	25	1.1	25	19	US-10-717-597-4354	Sequence 4354, Ap
43	25	1.1	25	19	US-10-717-597-4355	Sequence 4355, Ap
44	25	1.1	25	19	US-10-717-597-4356	Sequence 4356, Ap
45	25	1.1	25	19	US-10-717-597-4357	Sequence 4357, Ap

#### ALIGNMENTS

#### RESULT 1

US-10-131-827-743  
; Sequence 743, Application US/10131827  
; Publication No. US20040009479A1  
; GENERAL INFORMATION:  
; APPLICANT: Wohlgenuth, Jay  
; APPLICANT: Fry, Kirk  
; APPLICANT: Woodward, Robert  
; APPLICANT: Ly, Ngoc  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND MONITORING AUTOIMMUNE  
; FILE OF INVENTION: CHRONIC INFLAMMATORY DISEASES  
; CURRENT APPLICATION NUMBER: US/10/131,827  
; CURRENT FILING DATE: 2002-09-06  
; PRIOR APPLICATION NUMBER: US 10/006,290  
; PRIOR FILING DATE: 2001-10-22  
; PRIOR APPLICATION NUMBER: US 60/296,764  
; PRIOR FILING DATE: 2001-06-08  
; NUMBER OF SEQ ID NOS: 9090  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 743  
; LENGTH: 50  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-131-827-743

Query Match 2.2%; Score 50; DB 17; Length 50;  
Best Local Similarity 100.0%; Pred. No. 0.013;  
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;













Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	37.4	1.7	48	6	AR178317	Sequence
2	37.4	1.7	48	6	AX323399	Sequence
3	37.4	1.7	50	6	A51711	Sequence 17
4	37.4	1.7	50	6	AR167590	Sequence
5	37.4	1.7	50	6	AR178300	Sequence
6	37.4	1.7	50	6	AX323382	Sequence
7	37.4	1.7	50	6	AX323382	Sequence
8	35.2	1.6	48	6	AR199544	Sequence
9	35.2	1.6	48	6	AR374696	Sequence
10	35.2	1.6	48	6	AR409331	Sequence
11	35.2	1.6	48	6	AX239593	Sequence
12	35.2	1.6	48	6	BD056458	Novel 10w
C 13	31	1.4	39	6	I84408	Sequence 9
C 14	28.2	1.3	49	6	CO002787	Sequence
15	28	1.2	36	6	A91851	Sequence 10
16	28	1.2	45	6	AR199538	Sequence
17	28	1.2	45	6	AR374690	Sequence
18	28	1.2	45	6	AR409325	Sequence
19	28	1.2	45	6	AX239587	Sequence

```

/mmol_type="unassigned DNA"

Query Match      1.7%; Score 37.4; DB 6; Length 50;
Best Local Similarity 87.2%; Pred. No. 2e+04;
Matches 41; Conservative 0; Mismatches 6; Indels 0; Gap

Qy 231 GAGGACGAGGAGGAAGAAGAGGAGGAGGAAGAGGAGGACGAGGACGGA 277
      |||||
Db 4 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGA 50

RESULT 5
LOCUS AR178300 50 bp DNA linear PAT 20-A
DEFINITION Sequence 17 from patent US 6319672.
ACCESSION AR178300
VERSION AR178300.1 GI:20219438
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 50)
AUTHORS Crouzet,J., Scherman,D., Wils,P., Blanche,P. and Cameron,B.
TITLE Purification of a triple helix formation with an immobilized oligonucleotide
JOURNAL Patent: US 6319672-A 17 20-NOV-2001;
FEATURES
    source
        location/Qualifiers
            1..50
            /organism="unknown"
            /mol_type="unassigned DNA"

ORIGIN

Query Match      1.7%; Score 37.4; DB 6; Length 50;
Best Local Similarity 87.2%; Pred. No. 2e+04;
Matches 41; Conservative 0; Mismatches 6; Indels 0; Gap

Qy 231 GAGGACGAGGAGGAAGAAGAGGAGGAGGAGGAGGAGGACGAGGACGGA 277
      |||||
Db 4 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGA 50

RESULT 6
LOCUS AX323382 50 bp DNA linear PAT 07-J
DEFINITION Sequence 17 from Patent WO0192511.
ACCESSION AX323382
VERSION AX323382.1 GI:18094144
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
        other sequences; artificial sequences.
REFERENCE 1
AUTHORS Crouzet,J., Scherman,D., Wils,P., Blanche,P. and Cameron,B.
TITLE Purification of a triple helix formation with an immobilized oligonucleotide
JOURNAL Patent: WO 0192511-A 17 06-DEC-2001;
        Aventis Pharma (FR)
FEATURES
    source
        location/Qualifiers
            1..50
            /organism="synthetic construct"
            /mol_type="unassigned DNA"
            /db_xref="taxon:32630"
            /note="synthetic oligonucleotide"

ORIGIN

Query Match      1.7%; Score 37.4; DB 6; Length 50;
Best Local Similarity 87.2%; Pred. No. 2e+04;
Matches 41; Conservative 0; Mismatches 6; Indels 0; Gap

Qy 231 GAGGACGAGGAGGAAGAAGAGGAGGAGGAGGAGGAGGACGAGGACGGA 277
      |||||
Db 4 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGA 50

```

**SOURCE**  
**ORGANISM**



Qy 235 ACGAGGAGGAAGAAGAGGAGGAGGAGGAGGAG 270  
| | | | | | | | | | | | | | | | | | | | | |  
Db 1 AGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 36

Search completed: August 4, 2005, 00:26:44  
Job time : 9532 secs

THIS PAGE BLANK (USPTO)



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 3, 2005, 17:10:58 ; Search time 1182 Seconds  
(without alignments)

11218.461 Million cell updates/sec

Title: US-09-544-776-1

Perfect score: 2240

Sequence: 1 cgtcacacagtagtgcctc.....taaaaaaaaaaaaaaaaaaa 2240

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 4167226

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N\_Geneseq\_16Dec04:\*

1: Geneseqn1980s:\*

2: Geneseqn1990s:\*

3: Geneseqn2000s:\*

4: Geneseqn2001as:\*

5: Geneseqn2001bs:\*

6: Geneseqn2002as:\*

7: Geneseqn2002bs:\*

8: Geneseqn2003as:\*

9: Geneseqn2003bs:\*

10: Geneseqn2003cs:\*

11: Geneseqn2003ds:\*

12: Geneseqn2004as:\*

13: Geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	50	2.2	50	6 ABZ00752	Abz00752 Human leu
2	50	2.2	50	10 ADG33274	Adg33274 Human DNA
3	37.4	1.7	48	6 AAS21106	Aas21106 (GGA)16 D
4	37.4	1.7	50	2 AAT32776	Aat32776 Homopurin
5	37.4	1.7	50	6 AAS19342	Aas19342 Plasmid X
6	35.2	1.6	48	5 AAH26508	Aah26508 Low densi
c 7	34	1.5	42	4 AAF31000	Aaf31000 Human NOG
8	34	1.5	43	4 AAF30999	Aaf30999 Human NOG
9	32.4	1.4	42	10 ABZ22128	Abz22128 Polyanion
10	32.4	1.4	42	10 ABZ22092	Abz22092 Polyanion
c 11	28.2	1.3	49	4 AAL28219	Aal28219 Human SNP
12	28	1.2	45	5 AAH28502	Aah28502 Low densi
13	27.4	1.2	31	4 AAI29843	Aai29843 Human sin
c 14	27.4	1.2	42	4 AAI64451	Aai64451 SSR motif
15	26.6	1.2	42	4 AAS13776	Aas13776 Simple se
c 16	26.6	1.2	46	4 AAL28459	Aal28459 Human SNP
c 17	26.2	1.2	41	4 AAL28691	Aal28691 Human SNP
c 18	26	1.2	42	4 AAS13779	Aas13779 Simple se
c 19	26	1.2	50	10 ADC56858	Adc56858 Micro gol
c 20	26	1.2	50	10 ADC56859	Adc56859 Micro gol

c 21	25.8	1.2	30	10 ABZ22048	Abz22048 Polyanion
22	25.8	1.2	45	2 AAV55838	Aav55838 Sequence
c 23	25.8	1.2	50	3 AAC26760	Aac26760 Human sec
24	25.6	1.1	47	10 ADJ81490	Adj81490 Plant ret
25	25.6	1.1	50	2 AAV76464	Aav76464 Staphyloc
26	25.4	1.1	40	9 ABZ59596	Abz59596 Human par
c 27	25.2	1.1	30	3 ABN81202	Abn81202 Litopenae
c 28	25.2	1.1	30	10 ABX04111	Abx04111 Oligonuc
29	25.2	1.1	30	10 ABZ22047	Abz22047 Polyanion
30	25.2	1.1	30	10 ABZ22086	Abz22086 Polyanion
c 31	25.2	1.1	42	2 AAT17027	Aat17027 Human mit
32	25.2	1.1	50	6 ABK34441	Abk94441 Human BRC
c 33	25	1.1	25	3 AAC64407	Aac64407 Human NOG
34	25	1.1	25	12 ADP17627	Adp17627 Renal cel
35	25	1.1	25	12 ADP17614	Adp17614 Renal cel
36	25	1.1	25	12 ADP17617	Adp17617 Renal cel
37	25	1.1	25	12 ADP17618	Adp17618 Renal cel
38	25	1.1	25	12 ADP17619	Adp17619 Renal cel
39	25	1.1	25	12 ADP17613	Adp17613 Renal cel
40	25	1.1	25	12 ADP17616	Adp17616 Renal cel
41	25	1.1	25	12 ADP17622	Adp17622 Renal cel
42	25	1.1	25	12 ADP17623	Adp17623 Renal cel
43	25	1.1	25	12 ADP17628	Adp17628 Renal cel
44	25	1.1	25	12 ADP17820	Adp17820 Renal cel
45	25	1.1	25	12 ADP17624	Adp17624 Renal cel

#### ALIGNMENTS

#### RESULT 1

ABZ00752

ID ABZ00752 standard; DNA; 50 BP.

AC ABZ00752;

XX

DT 09-JAN-2003 (first entry)

XX

DE Human leukocyte gene expression profiling probe SEQ ID NO 743.

XX

KW T7: leukocyte; gene expression profiling; allograft rejection;

KW atherosclerosis; congestive heart failure; systemic lupus erythematosus;

KW rheumatoid arthritis; osteoarthritis; cytomegalovirus; infection; probe;

XX

OS Homo sapiens.

XX

XX

PN WO200257414-A2.

XX

PD 25-JUL-2002.

XX

PF 22-OCT-2001; 2001WO-US047856.

XX

PR 20-OCT-2000; 2000US-0241994P.

XX

PA 08-JUN-2001; 2001US-0296764P.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

CC for leukocyte expression profiling. It is particularly useful for  
 CC diagnosing a disease, monitoring (rate of) progression of a disease,  
 CC predicting therapeutic outcomes, determining prognosis for a patient,  
 CC predicting disease complications in an individual or monitoring response  
 CC to treatment in an individual. The diseases include cardiac allograft  
 CC rejection, kidney allograft rejection, liver allograft rejection,  
 CC atherosclerosis, congestive heart failure, systemic lupus erythematosus,  
 CC rheumatoid arthritis, osteoarthritis or cytomegalovirus infection  
 XX

XX Sequence 50 BP; 13 A; 6 C; 13 G; 18 T; 0 U; 0 Other;  
 SQ Query Match 2.2%; Score 50; DB 6; Length 50;  
 Best Local Similarity 100.0%; Pred. No. 0.8;  
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2140 TGTGGTTTAAGCTGTACTGAACCTAAATCTGTGGAATGCATTGTGAACCTGT 2189  
 |||||  
 Db 1 TGTGGTTTAAGCTGTACTGAACCTAAATCTGTGGAATGCATTGTGAACCTGT 50

RESULT 2  
 ADG33274  
 ID ADG33274 standard; DNA; 50 BP.  
 XX AC  
 AC ADG33274;  
 XX  
 XX 26-FEB-2004 (first entry)  
 XX

DE Human DNA probe used to monitor expression of diagnostic genes SeqID598.  
 XX human; ss; autoimmune; chronic inflammatory disease; SLE;  
 KW systemic lupus erythematosus; rheumatoid arthritis; cholecystitis;  
 KW Sjogren's disease; CREST syndrome; scleroderma; ankylosing spondylitis;  
 KW ulcerative colitis; primary sclerosing cholangitis; appendicitis;  
 KW diverticulitis; primary biliary sclerosis; probe.  
 XX

OS Homo sapiens.  
 XX  
 XX WO2003090694-A2.  
 PN  
 XX 06-NOV-2003.  
 PD  
 XX 24-APR-2003; 2003WO-US013015.  
 XX  
 XX 24-APR-2002; 2002US-00131827.  
 PR  
 XX (EXPR-) EXPRESSION DIAGNOSTICS INC.  
 PA  
 XX Wohlgemuth J, Fry K, Woodward R, Ly N;  
 PI  
 XX WPI; 2003-877243/81.  
 DR  
 XX Diagnosing or monitoring autoimmune and chronic inflammatory diseases,  
 PT such as rheumatoid arthritis, systemic lupus erythematosus, ulcerative  
 PT colitis, psoriasis and asthma by detecting the expression level of one or  
 PT more genes.  
 XX

PS Claim 1; SEQ ID NO 598; 877pp; English.  
 XX

CC This invention relates to novel methods for diagnosing and monitoring  
 CC autoimmune and chronic inflammatory diseases. Specifically, it refers to  
 CC the identification of genes that have a clinical utility as diagnostic  
 CC tools for the management of, in particular, patients with systemic lupus  
 CC erythematosus (SLE) or rheumatoid arthritis (RA). Accordingly, the  
 CC present invention describes a method for determining the levels of  
 CC multiple differentially expressed genes of a patient, in a concerted  
 CC manner, in order to achieve an improved diagnostic assay with sensitivity  
 CC and specificity for the disease in question. As such, these genes are  
 CC useful for the diagnosis of various other inflammatory disorders  
 CC including cholecystitis, Sjogren's disease, CREST syndrome, scleroderma,  
 CC ankylosing spondylitis, ulcerative colitis, primary sclerosing  
 CC cholangitis, appendicitis, diverticulitis, and primary biliary sclerosis.  
 CC This oligonucleotide is a human DNA probe used to monitor the expression

CC level of the differentially expressed diagnostic genes of the invention.  
 XX

SQ Sequence 50 BP; 13 A; 6 C; 13 G; 18 T; 0 U; 0 Other;  
 Query Match 2.2%; Score 50; DB 10; Length 50;  
 Best Local Similarity 100.0%; Pred. No. 0.8;  
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2140 TGTGGTTTAAGCTGTACTGAACCTAAATCTGTGGAATGCATTGTGAACCTGT 2189  
 |||||  
 Db 1 TGTGGTTTAAGCTGTACTGAACCTAAATCTGTGGAATGCATTGTGAACCTGT 50

RESULT 3  
 AAS21106  
 ID AAS21106 standard; DNA; 48 BP.  
 XX AC  
 AC AAS21106;  
 XX  
 XX 20-MAR-2002 (first entry)  
 XX

DE (GGA)16 DNA purification oligonucleotide.  
 XX ss; DNA purification; triple helix; plasmid purification.  
 KW  
 XX Synthetic.  
 OS

FH Key Location/Qualifiers  
 FT repeat\_region 1..48  
 FT /\*tag= a  
 FT /rpt\_type= TANDEM  
 FT repeat\_unit 1..3  
 FT /\*tag= b  
 FT /note= "GAA repeat type"  
 XX

PN WO200192511-A2.  
 XX  
 XX 06-DEC-2001.  
 PD  
 XX 25-MAY-2001; 2001WO-US017122.  
 PF  
 XX 26-MAY-2000; 2000US-00580923.  
 PR  
 XX (AVET ) AVENTIS PHARMA SA.  
 PA  
 XX Crouzet J, Scherman D, Wills P, Blanche F, Cameron B;  
 PI  
 XX WPI; 2002-097772/13.  
 DR

PT Purifying double-stranded (ds) DNA from a solution containing dsDNA and  
 PT other components, comprises passing the solution through a support  
 PT comprising a covalently coupled oligonucleotide able to form a triple  
 PT helix with the dsDNA.  
 XX

XX Example 7; Page 20; 40pp; English.  
 PS

CC This invention comprises a method of purifying double-stranded DNA from a  
 CC solution containing the double-stranded DNA mixed with other components,  
 CC comprising passing the solution through a support comprising a covalently  
 CC coupled oligonucleotide capable of forming a triple helix with the double  
 CC -stranded DNA by hybridisation with a specific sequence present in the  
 CC double-stranded DNA. The method is useful for purifying double-stranded  
 CC DNA contained in a solution and mixed with other components. The new  
 CC method is a simple, rapid and effective method for DNA purification, and  
 CC makes it possible to obtain especially high purities with high yields.  
 CC The method enables DNA to be purified from complex mixtures comprising  
 CC other nucleic acids, proteins, endotoxins, nucleases and the like. The  
 CC supports may be readily recycled, and the DNAs obtained display improved  
 CC properties to pharmaceutical safety. Further, the method entails only one  
 CC step contrary to prior art. The present sequence represents a DNA  
 CC sequence contained within the plasmid pXL2725. This sequence is used for  
 CC purification of this plasmid using the method of the invention  
 XX

SQ		Sequence 48 BP; 16 A; 0 C; 32 G; 0 T; 0 U; 0 Other;	
	Query Match	1.7%; Score 37.4; DB 6; Length 48;	
	Best Local Similarity	87.2%; Pred. No. 3.4e+02;	
	Matches	41; Conservative 0; Mismatches 6; Indels 0; Gaps 0;	
QY		231 GAGGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGCA 277 	
Dd		2 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGA 48 	
	RESULT 4		
	AAT32776		
ID		AAT32776 standard; DNA; 50 BP.	
XX			
AC		AAT32776;	
XX			
DT		18-FEB-1997 (first entry)	
XX			
DE		Homopurine target for triple helix-forming oligonucleotide.	
XX			
KW		Triple helix; triplex formation; Hoogsteen base pairing; plasmid;	
KW		purification; double-stranded DNA; homopyrimidine; polypurine; ss.	
XX			
OS		Synthetic.	
XX			
PV		WO9618744-A2.	
XX			
PD		20-JUN-1996.	
XX			
PF		08-NOV-1995; 95WO-FR001468.	
XX			
PR		16-DEC-1994; 94FR-00015162.	
XX			
PA		(RHON ) RHONE POULENC RORER SA.	
XX			
PI		Crouzet J, Scherman D, Wils P;	
DR		WPI; 1996-300660/30.	
XX			
PT		Purificn. of double stranded DNA by triple helix formation - comprises hybridising immobilised oligo-nucleotide to specific sequence in target DNA.	
PS		Claim 12; Page 25; 34pp; French.	
CC		Double-stranded (ds) DNA can be purified from complex mixtures of nucleic acids, proteins, endotoxins, nucleases, etc. by passing the mixture over a support to which an oligonucleotide is covalently attached; the oligonucleotide is able to form a triple helix by hybridisation with a specific sequence present in the dsDNA. The present sequence is a preferred homopurine target for triplex formation with a single-stranded oligonucleotide. The target sequence may be present naturally, e.g. in a plasmid origin of replication, or can be introduced artificially. The method is particularly suited to purification of plasmid DNA	
XX			
SQ		Sequence 50 BP; 17 A; 0 C; 33 G; 0 T; 0 U; 0 Other;	
	Query Match	1.7%; Score 37.4; DB 2; Length 50;	
	Best Local Similarity	87.2%; Pred. No. 3.4e+02;	
	Matches	41; Conservative 0; Mismatches 6; Indels 0; Gaps 0;	
QY		231 GAGGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGCA 277 	
Dd		4 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGA 50 	
	RESULT 5		
	AAS19342		
ID		AAS19342 standard; DNA; 50 BP.	
XX			
AC		AAS19342;	
XX			
DT		12-NOV-2001 (first entry)	
XX			

DT	20-MAR-2002 (first entry)	
XX		
DE	Plasmid XL2725 sequence.	
XX		
KW	ds; DNA purification; triple helix; plasmid purification; XL27256.	
XX		
OS	Synthetic.	
XX		
FH	Key	Location/Qualifiers
FT	repeat_region	6..50
FT		/tag= a
FT		/rpt_type= TANDEM
FT	repeat_unit	6..8
FT		/tag= b
FT		/notes= "GGA repeat type"
XX		
PN	WO200192511-A2.	
XX		
PD	06-DEC-2001.	
XX		
PF	25-MAY-2001; 2001WO-US017122.	
XX		
PR	26-MAY-2000; 2000US-00580923.	
XX		
PA	(AVET ) AVENTIS PHARMA SA.	
XX		
PI	Crouzet J, Scherman D, Wils P, Blanche F, Cameron B;	
XX		
DR	WPI; 2002-097772/13.	
XX		
PT	Purifying double-stranded (ds) DNA from a solution containing dsDNA and other components, comprises passing the solution through a support comprising a covalently coupled oligonucleotide able to form a triple helix with the dsDNA.	
XX		
PS	Example 7.1; Page 20; 40pp; English.	
XX		
CC	This invention comprises a method of purifying double-stranded DNA from a solution containing the double-stranded DNA mixed with other components, comprising passing the solution through a support comprising a covalently coupled oligonucleotide capable of forming a triple helix with the double-stranded DNA by hybridisation with a specific sequence present in the double-stranded DNA. The method is useful for purifying double-stranded DNA contained in a solution and mixed with other components. The new method is a simple, rapid and effective method for DNA purification, and makes it possible to obtain especially high purities with high yields. The method enables DNA to be purified from complex mixtures comprising other nucleic acids, proteins, endotoxins, nucleases and the like. The supports may be readily recycled, and the DNAs obtained display improved properties to pharmaceutical safety. Further, the method entails only one step contrary to prior art. The present sequence represents the sequence of the plasmid XL2725 verified upon sequencing, this sequence differs from the expected sequence (GGA)17	
XX		
SQ	Sequence 50 BP; 17 A; 0 C; 33 G; 0 T; 0 U; 0 Other;	
	Query Match	1.7%; Score 37.4; DB 6; Length 50;
	Best Local Similarity	87.2%; Pred. No. 3.4e+02;
	Matches	41; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY		231 GAGGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGCA 277 
Dd		4 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGA 50 
	RESULT 6	
	AAH26508	
ID		AAH26508 standard; DNA; 48 BP.
XX		
AC		AAH26508;
XX		
DT		12-NOV-2001 (first entry)
XX		

	SQ	Sequence 48 BP; 16 A; 0 C; 32 G; 0 T; 0 U; 0 Other;	
		Query Match Best Local Similarity 1.7%; Score 37.4; DB 6; Length 48; Matches 41; Conservative 0; Mismatches 6; Indels 0; Gaps 0;	
QY	Db	231 GAGGACGGAGGAGNAGAAGAGGAGGAGGAGGAGGAGGAGGAGCAGCA 277   2 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGA 48	
		RESULT 4 AAT32776 ID AAT32776 standard; DNA; 50 BP. XX AC AAT32776; XX DT 18-FEB-1997 (first entry) XX DE Homopurine target for triple helix-forming oligonucleotide.	
		DE Triple helix; triplex formation; Hoogsteen base pairing; plasmid; KW purification; double-stranded DNA; homopyrimidine; polypurine; ss. OS Synthetic. XX WO9618744-A2. PN XX PD 20-JUN-1996. XX PF 08-NOV-1995; 95WO-FR001468. XX PR 16-DEC-1994; 94FR-00015162. PA (RHON ) RHONE POULENC RORER SA. XX PI Crouzet J, Scherman D, Wils P; PI WPI; 1996-300660/30. DR Purificn. of double stranded DNA by triple helix formation - comprises hybridising immobilised oligo-nucleotide to specific sequence in target PT DNA. PT Claim 12; Page 25; 34pp; French. PS Double-stranded (ds) DNA can be purified from complex mixtures of nucleic acids, proteins, endotoxins, nucleases, etc. by passing the mixture over CC a support to which an oligonucleotide is covalently attached; the CC oligonucleotide is able to form a triple helix by hybridisation with a CC specific sequence present in the dsDNA. The present sequence is a CC preferred homopurine target for triplex formation with a single-stranded CC oligonucleotide. The target sequence may be present naturally, e.g. in a CC plasmid origin of replication, or can be introduced artificially. The CC method is particularly suited to purification of plasmid DNA XX SQ Sequence 50 BP; 17 A; 0 C; 33 G; 0 T; 0 U; 0 Other;	
		Query Match Best Local Similarity 1.7%; Score 37.4; DB 2; Length 50; Matches 41; Conservative 0; Mismatches 6; Indels 0; Gaps 0;	
QY	Db	231 GAGGACGGAGGAGNAGAAGAGGAGGAGGAGGAGGAGGAGGAGCAGCA 277   4 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGA 50	
		RESULT 5 AAS19342 ID AAS19342 standard; DNA; 50 BP. XX AC AAS19342; XX DT 12-NOV-2001 (first entry)	
		DT 20-MAR-2002 (first entry) XX Plasmid XL2725 sequence. XX ds; DNA purification; triple helix; plasmid purification; XL27256. XX Synthetic. XX Key repeat_region Location/Qualifiers FT FT /tag= a FT FT /rpt_type= TANDEM FT FT .8 FT FT /*tag= b FT FT /notes= "GGA repeat type" XX WO200192511-A2. XX PD 06-DEC-2001. XX PF 25-MAY-2001; 2001WO-US017122. XX PR 26-MAY-2000; 2000US-00580923. XX PA (AVET ) AVENTIS PHARMA SA. XX PI Crouzet J, Scherman D, Wils F, Blanche F, Cameron B; XX WPI; 2002-097772/13. XX Purifying double-stranded (ds) DNA from a solution containing dsDNA and other components, comprises passing the solution through a support comprising a covalently coupled oligonucleotide able to form a triple helix with the dsDNA. XX Example 7.1; Page 20; 40pp; English. XX This invention comprises a method of purifying double-stranded DNA from a solution containing the double-stranded DNA mixed with other components, comprising passing the solution through a support comprising a covalently coupled oligonucleotide capable of forming a triple helix with the double-stranded DNA by hybridisation with a specific sequence present in the double-stranded DNA. The method is useful for purifying double-stranded DNA contained in a solution and mixed with other components. The new method is a simple, rapid and effective method for DNA purification, and makes it possible to obtain especially high purities with high yields. The method enables DNA to be purified from complex mixtures comprising other nucleic acids, proteins, endotoxins, nucleases and the like. The supports may be readily recycled, and the DNAs obtained display improved properties to pharmaceutical safety. Further, the method entails only one step contrary to prior art. The present sequence represents the sequence of the plasmid XL2725 verified upon sequencing, this sequence differs from the expected sequence (GGA)17 XX SQ Sequence 50 BP; 17 A; 0 C; 33 G; 0 T; 0 U; 0 Other;	
		Query Match Best Local Similarity 1.7%; Score 37.4; DB 6; Length 50; Matches 41; Conservative 0; Mismatches 6; Indels 0; Gaps 0;	
QY	Db	231 GAGGACGGAGGAGNAGAAGAGGAGGAGGAGGAGGAGGAGGAGCAGCA 277   4 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGA 50	
		RESULT 6 AAH26508 ID AAH26508 standard; DNA; 48 BP. XX AC AAH26508; XX DT 12-NOV-2001 (first entry)	

[illegible]







[illegible]

